DOCUMENT RESUME

EA 001 669

ED 023 176

By-Klausmeier, Herbert J., Ed. And Others
Research and Development Activities in R & I Units of Five Elementary Schools of Racine, Wisconsin, 1966-67. Report From Project Models.
Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
Spons Agency-Office of Education (DHEW), Washington, DC. Bureau of Research.
Report No-TR-52
Bureau No-BR-5-0216
Pub Date Apr 68
Contract-OEC-5-10-154
Note-53p.
EDRS Price MF-$0.25 HC-$2.75
Descriptors-Academic Achevement, *Culturally Disadvantaged, *Elementary Schools, Field Studies, Handwriting Instruction, *Individualized Programs, Instructional Improvement, Language Arts, *Learning Motivation, Mathematics Instruction, Parent Attitudes, Parent School Relationship, *Research and Instruction Units, Spelling Instruction, Statistical Analysis, Teacher Attitudes
Identifiers-Project MODELS

Experimental educational programs emphasizing individualized instruction and motivation were conducted with classes from kindergarten through grade 6. Projects analyzed individualization and motivational procedures in mathematics, individualization in handwriting and spelling, instructional procedures in language arts for disadvantaged children, and the effects of increased home-school contact on parental attitudes and student achievement. Field testing revealed which control groups for the Research and Instruction Units were adequate and also which strategies for ascertaining pupil achievements were appropriate. Administration of a teacher opinion scale confirmed that student behavior, achievement, motivation, and attitudes were better in R & I classes than in control classrooms. (JK)

BR-5-0216
PA-24

RESEARCH AND DEVELOPMENT ACTIVITIES IN R&I UNITS OF FIVE ELEMENTARY SCHOOLS OF RACINE, WISCONSIN, 1966-1967

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR
COGNITIVE LEARNING

EA 001 669

Technical Report No. 52

RESEARCH AND DEVELOPMENT ACTIVITIES IN R & I UNITS OF FIVE ELEMENTARY SCHOOLS OF RACINE, WISCONSIN, 1966–1967

Herbert J. Klausmeier, Mary Quilling,
and James L. Wardrop, Editors

Report from Project MODELS
Richard G. Morrow and Herbert J. Klausmeier, Principal Investigators

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

April 1968

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program.

Center No. C-03 / Contract OE 5-10-154

PREFACE

A major objective of the Wisconsin Research and Development Center for Cognitive Learning is to develop an environment in local school buildings and systems which facilitates student learning and also research, development, and innovative activities. This report is concerned with the description and evaluation of such facilitative organizations and their activities in five elementary schools in the Racine Unified School District. The report further demonstrates how instructional and supervisory personnel in the public schools, working with personnel at the Center who possess specialized knowledge in various disciplines, cooperate to extend knowledge and improve educational practice through research and development.

Many people other than the R & D personnel and Unit leaders denoted as authors contributed their skills in planning, executing, or evaluating the activities reported herein. In the Racine Schools Mr. John LeBlanc, Mathematics Consultant, Miss Elizabeth Williams and Mr. Neil Vail, Language-Arts Consultants, and Miss Mildred Brady, Reading Consultant, aided in the planning of the experiments. Mr. Harris Russell, Director of Instructional Services, planned with Dr. Glenn Tagatz and Dr. James Wardrop, postdoctoral fellows at the Center, the field testing program.

Professor Herbert J. Klausmeier, Principal Investigator of Project MODELS and Director of the Wisconsin R & D Center, initiated the idea of R & I Units and assumed primary responsibility for the conceptualization of the total R & I program and for the broad implementation strategies in the local schools. Professor Klausmeier wrote the introductory and concluding sections of this report, and Mrs. Mary Quilling edited it. Mrs. Doris Cook assumed primary responsibility for working with the building personnel during the year. She, Mr. Tom Houston, Miss Thelma Baldwin, Dr. Tagatz, and Dr. Wardrop served as consultants for the experiments, and Dr. Wardrop assumed primary responsibility for the field-testing report. Other Center personnel who assisted in data collection and analyses include Mrs. Barbara Kennedy and Mr. James Bavry. The editors and authors acknowledge with appreciation the contributions of the above.

Thomas A. Romberg
Director, Program 3

CONTENTS

		Page
	List of Tables	vii
	List of Figures	ix
	Abstract	xi
I	Introduction	1
II	Individualization and Motivational Procedures in Mathematics	3
	Howell School, Second-Grade Unit	3
	Stephen Bull Primary Unit	4
	Stephen Bull Third-Grade Unit	7
	Discussion of Experiments in Individualization and Motivation of Arithmetic Instruction	9
III	Individualization in Handwriting and Spelling, Giese School	11
	Third-Grade Unit	11
	The Comparative Effects of Three Methods of Handwriting Instruction upon Pupil Skill in Handwriting	12
	Fourth-Grade Unit	16
	The Comparative Effects of Three Methods of Handwriting Instruction	17
IV	Instructional Procedures in Language Arts for Disadvantaged Children	21
	Written Expression, Winslow School, Fifth Grade	21
	The Effects of Motivational Procedures on Children's Reading, Stephen Bull, Third Grade	22
	Improving Oral and Written Expression, Franklin School, Second Grade	24
	Language Arts for Disadvantaged Children, Franklin School, Kindergarten	25
V	The Effects of Increased Home-School Contact on Parent Attitudes and Student Achievement, Howell School	30
VI	Field Testing R & I Units, Racine Unified School District 1966–67	33
	The Second-Grade Unit at Howell School	33
	The Fourth-Grade Unit at Giese School	34
	The Fifth-Grade Unit at Howell School	35

		Page
	The Fifth-Grade Unit at Winslow School	37
	General Results of Exploratory Field Testing: Problems and Recommendations	39
VII	Teacher Opinion Scale	41
VIII	Concluding Statement	43

LIST OF TABLES

Table		Page
1	Classification of Subjects	4
2	Tests of Significance Using Wilks' Lambda Criterion	5
3	Means and Standard Deviations for the Two Treatment Groups	5
4	Means for Each Sex and Age Group	5
5	Pretest and Posttest Means for First Year Primary Pupils on Teacher-Constructed Test	6
6	Analysis of Covariance on Posttest Total Scores of Second Year Primary Pupils	6
7	Pretest and Posttest Means for Second Year Primary Pupils, Teacher-Constructed Test	7
8	Second-Year Primary Students' Mean Performance on Kuhlmann-Anderson Intelligence Test and on Stanford Achievement Tests, Primary II Battery, Arithmetic Subtests	7
9	Distribution of Children in Experimental and Control Groups	8
10	Mean Totals for Subgroups on Teacher-Constructed Test	8
11	Analysis of Covariance on Teacher-Constructed Posttest	9
12	Adjusted Means of Subclasses	9
13	Experimental and Control Treatments in Three Arithmetic Experiments	9
14	Distribution of Children in Experimental and Control Groups, Giese School, Grade 3	13
15	Summary of Analyses of Covariance, Grade 3	14
16	Means and Standard Deviations for Significant "Treatment" Effects, Grade 3	14
17	Unadjusted Means for Treatment x Sex Interaction, Number of Words Fastest Measure	15

Table		Page
18	Normal Condition Ratings for Giese Third Graders and a Statewide Sample of Fourth Graders	15
19	Distribution of Children in Experimental and Control Groups, Giese School, Grade 4	18
20	Mean Squares and F Ratios for Speed and Legibility of Handwriting, Grade 4	19
21	Means for Significant "Sex" Effects	19
22	Means for Significant "Ability Level" Effects for Sample from Other Area	19
23	Cell Means for Significant Treatment x Sex Interaction for Number of Words Normal	20
24	Normal Condition Ratings for Giese Fourth Graders and a Statewide Sample of Fourth Graders	20
25	Stanford Achievement Test Mean Grade Equivalents Fall and Spring, Winslow School, Grade 5	22
26	Group Means for Measures Used in this Project	25
27	Mean Ratings for Oral Presentations	25
28	Design of Language Enrichment Experience, Franklin School	27
29	Summary of Analysis of Variance	27
30	Means for "Groups" Factor	27
31	Means for "Ability" Factor	28
32	Means for Groups x Measures Interaction	28
33	Means for Ability x Measures Interaction	28
34	Return Rate of Parent Questionnaires, Howell School, Grade 5	31
35	Analysis of Covariance Performed on Parent Questionnaires	31
36	Locations, Levels, and Instruments for Field Testing in Racine	34
37	Summary of Analyses of Covariance: Howell School, Grade 2	36
38	Howell and Control Schools, Grade 2, Fall and Spring Group Means on Stanford Achievement Test	36
39	Giese School, Grade 4, Fall and Spring, Group Means on Stanford Achievement Test	37
40	Summary of Analyses of Covariance: Howell School, Grade 5	38

Table		Page
41	Stanford Achievement Test Scores for Howell Grade 5 and Control School, Fall 1966 and Spring 1967	39
42	Means for Groups on Baseline Measures	39
43	Posttest Means and Gains on Stanford Achievement Subtests: Winslow School, Grade 5	39
44	Mean subscores on Teacher Opinion Questionnaire	41

LIST OF FIGURES

Figure		Page
1	Performance of Treatment Groups on Teacher-Constructed Test at Midterm and Conclusion of Experiment, Franklin School Kindergarten	29
2	Means of Treatment Groups on Fall and Spring Administration of Attitude Questionnaire	31
3	Profile of Means on Stanford Achievement Test for Experimental and Control Groups, Grade 5, Howell School, Spring 1967	34
4	Mean Performance of Howell and Control School Pupils, Stanford Achievement Test, Fall 1966	35
5	Mean Performance of Giese and Control School Pupils, Stanford Achievement Test, Fall 1966	37

ABSTRACT

This report summarizes the activities of the R & I (Research and Instruction) Units in five elementary schools of the Racine Unified School District during the 1966-67 school year. Among the significant findings of the several experiments performed are the following:

- A well planned language enrichment program for disadvantaged kindergarten children improves their communication skills.
- Written expression and verbal achievement of intermediate pupils is effectively stimulated by means of cartoons.
- Pupils who are taught handwriting by individualized methods outperform those taught by traditional programs.
- Extensive home-school contacts result in improved parental attitude toward the school and better school achievement of the students.
- The informative feedback provided by both a teacher-pupil conference and a mastery checklist results in superior performance in arithmetic.
- Concrete rewards and individual conferences are together effective in stimulating disadvantaged children's reading of library books.

1

INTRODUCTION

Securing more efficient pupil learning continues to be the main focus of the research and development activities conducted jointly by the R & D Center and several school systems as part of Project MODELS. One possible means for accomplishing this is to replace the graded, self-contained classroom with a research and instruction unit (R & I Unit) in which various instructional activities may be performed more effectively. R & I Units were organized in five elementary schools of Racine during 1966–1967. In each Unit the attempt was made (1) to provide excellent instruction for children, (2) to carry out research which is essential for improving instruction, (3) to develop new instructional procedures, materials, or ideas for improving instruction, and (4) to bring into the Unit promising educational innovations. The R & I Units are hypothesized to be more effective than self-contained classrooms in achieving these purposes. In order to be more effective, the roles of the building principal, Unit leader, classroom teacher, and teaching aide are being refined, and new relationships involving representatives of the central staff, the school building, and other agencies are being established. Thus, the concept of improving instruction through research and development in R & I Units is complex, involving an attempt to utilize time, space, equipment, supplies, instructional methods, instructional personnel, subject-matter content and sequence, and evaluation procedures in a more effective manner to achieve an efficient total educational program for each child.

When dealing with a total program, more time is required to get the various components integrated. However, the possibility for making significant improvements is also large. During the first year, the major effort is necessarily upon achieving a smoothly operating instructional unit and gaining familiarity with research, development, and innovative procedures. While this is being done, large gains in student learning should not be expected. Once the instructional staff and children operate as a unit and better materials and methods are developed, research, and utilized, we may anticipate substantial improvement in student learning.

The two main instructional phenomena dealt with in the Units centered on individualizing instruction and motivation. Consultants from the R & D Center worked with the staff of the schools. Subject-matter consultants from the R & D Center or the central staff of the local school participated in decision-making where subject-matter specialization was called for in connection with the program of individualization.

The approach to individualization employed in the R & D Center is one of arranging a program of instruction for each child that will meet the various objectives of the educational program. This, in turn, calls for some instruction on a one-to-one basis, some small-group, and some large-group instruction.

In instruction on a one-to-one basis, the child proceeds at a rate appropriate for him. This type of individualized work with the teacher and independent study are required to meet those objectives concerned with the acquisition of independent skills. Some educational objectives require instruction in small groups. Pupils may be brought together in groups of 2 to 15 or more to work on specific activities of a fairly homogeneous type; for example, 2 to 15 children from a total group of 100 may be brought together for specific instruction related to acquisition of certain concepts or processes in arithmetic. Small groups also may be brought together to deal with the same word recognition skills. Small groups may be formed on the basis of interest, friendship, neighborhood, residence, and the like in social studies in connection with achieving certain objectives related to communication skills and attitude development. The extent to which large groups of 75 to 150 children may be brought together effectively has not been tested systematically. It is known that students may engage in individual study activities simultaneously in large groups. In the Units in the

elementary school, the principal reason for bringing all the students within the Unit together into the same group for part of the instructional day is to achieve better utilization of teacher time. Children participating in independent study or some other large group activity can proceed without all of the instructional staff of the Unit being present. This in turn frees part of the instructional staff during that period of time for planning, conferring, and executing other activities essential for making the small-group and one-to-one instructional activities work effectively.

Attention was also given throughout the year to research and development regarding motivation. Getting a larger number of students to want to learn and also to behave well is a continuing responsibility of R & I Units. We appear to have sufficient knowledge about the means of controlling behavior of young children so that few discipline problems should emerge in the elementary school. Devising procedures for applying this knowledge and testing some of the procedures is a continuing activity in R & I Units. From the preceding it may be properly inferred that no systematic attempt was made to improve instruction in any one subject-matter field in each Unit. This will be done more systematically during 1967-68.

In early 1966-67, a plan for field testing the Units, developed by Wardrop and Tegata (see p. 32), was reported in Working Paper No. 4 of the Center. Only parts of the total plan for field testing were executed during the 1966-67 school year. Also, the attempt was made to utilize the local resources of each school system in the field testing, including each school's testing program; therefore, the amount of information obtained regarding the Units varied within a school system and across school systems. In some of the elementary R & I Units field test data were gathered dealing with pupil achievement in the Units as measured by standardized tests. Instruments were developed and tested to secure opinions of pupils regarding the Units, and also the opinions of teachers and principals as to how well the research, development, and innovation functions were being achieved. In the main, then, field testing procedures and instruments were tried during the year, and the data obtained yield some preliminary information about the functioning of Units in settings of one or two Units in new elementary school buildings.

One of three instruments which was developed was an opinion scale to secure opinions of teachers and building principals regarding the instructional research, development, and innovation functions of the Unit and also to determine the effects of the Unit organization on teachers, students, and instructional practices. In the same opinionnaire information was secured regarding the utilization of resource persons and knowledge of individual students by the instructional staff. Another instrument, a check list, was developed for the purpose of determining the adequacy of the facilities, equipment, and supplies with respect to accomplishing the objectives of the R & I Units. A pupil questionnaire was developed for the purpose of securing information about the child, home, and neighborhood background, and opinions about a variety of matters associated with schooling. Finally, the field testing provided information regarding adequate control groups for R & I Units and also the extent to which different strategies for ascertaining pupil achievements were appropriate for this type of field study.

II

INDIVIDUALIZATION AND MOTIVATIONAL PROCEDURES IN MATHEMATICS

H. J. Klausmeier, Mary Quilling, Doris M. Cook,
James L. Wardrop, and John LeBlanc, R & D Center
Mary Jane Clausen, Pat Hansen, and Mae Elsdon,
Learning Specialists

Techniques of individualizing mathematics instruction and of motivating pupils were evaluated, under the guidance of John LeBlanc, mathematics consultant, in experiments in the Howell second-grade Unit, the Stephen Bull primary Unit, and the Stephen Bull third-grade Unit. All concepts and subconcepts in mathematics were identified for Grades 1, 2, and 3. Diagnostic tests were developed for each main concept and subconcept. These tests were used first to identify each child's standing related to the concept and subsequently to assess his progress in acquiring further knowledge and skills. Individual record folders listing all of the subconcepts in the form of behaviors as related to each more inclusive concept were used in charting each child's progress.

HOWELL SCHOOL, SECOND-GRADE UNIT

The R & I Unit at the second grade in Howell School, Mr. Glynn Humphrey, Principal, was organized for the first time in the fall of 1967. The Unit leader, Mary Jane Clausen, after attending the eight-week seminar on campus, attempted to develop an excellent instructional program for these 72 students, many of whom were from the Inner City. The teachers of this Unit were Annette Barnes, Martha Hutsick, and Hazel Palmer, certified teachers. The team met every Thursday for 1-1/2 hours and more frequently as the need arose.

Much additional help in reading was provided through the additional staff members of the Unit. One group who had been taught reading with the Initial Teaching Alphabet remained intact for reading instruction. A team approach to instruction was used for other students. The teachers teamed in almost all subject areas.

A highlight of this Unit's activities was a joint project in which fifth- and sixth-grade children served as models or helpers to whom the younger children could dictate stories. The stories were based on actual or imaginary field trips. This cooperative venture was beneficial for both groups of students. The model wrote the story as dictated, looked up unfamiliar words in the dictionary, and then rewrote the younger child's story. The stories were then typed and compiled into a newsletter which the second graders took home.

The children in this Unit also participated regularly in song fests which the Unit leader planned. These were held in the school gym for the entire Unit and appeared to be especially satisfying and stimulating to the entire group.

The Unit's major research and development activities, with which Dr. LeBlanc assisted, dealt with mathematics. The Unit staff assisted in identifying concepts and subconcepts, and in developing diagnostic tests. An experiment was conducted as a small part of the total program to ascertain the effect of weekly oral discussion of progress with each child.

Subjects

In the second grade at Howell School, a stratified random sample of 24 students was drawn from the entire second grade of 72 students to form the experimental group. This group was composed of 6 older girls, 6 older boys, 6 younger girls, and 6 younger boys.

Treatments

Standardized tests given to all pupils included the Stanford Achievement Test, in October, 1966, and the Kuhlmann-Anderson

Mental Abilities Test administered in November, 1966. Additionally, diagnostic tests for each of the major skill areas were given to each child prior to the study.

All pupils kept individual folders. Initially, the child's accomplishments to date were recorded. Then, as a child acquired a concept, on the basis of a diagnostic test, he was allowed to color in the squares related to this concept in his folder. All students used the Scott Foresman Seeing Through Arithmetic, Grade 2, and the SRA program for Grade 2.

Each child in the experimental group additionally had a weekly five-minute conference with the teacher. The child's progress in arithmetic was discussed through reference to his folder. At this time the child colored in the squares denoting his recent accomplishments. While the children in the control group also colored in squares, they did not have an individual conference nor did they refer to their folders weekly.

The 24 experimental subjects were randomly assigned to four teachers. The teachers rotated monthly among the four groups thus formed. The duration of the experiment was twenty weeks.

Data Collected

A teacher-constructed test covering concepts in the folder was used as a postexperimental measure. A second indicator of progress was the number of squares each child had colored in his folder during the experiment.

Analysis of the Data

Complete data at the conclusion of the experiment were available for 60 students, classified by sex, age, and treatment group (Table 1).

Table 1

Classification of Subjects

Sex	Age	Treatment: Experimental	Treatment: Control	Total
Male	Younger	5	12	17
	Older	4	10	14
Female	Younger	6	6	12
	Older	6	11	17
Total		21	39	60

Because there were two dependent variables, a multivariate analysis of variance was performed. The results of significance tests, based on Wilks' lambda criterion, are presented in Table 2.

The effect attributable to treatment differences is significant at the .10 level and favors the experimental group. The means and standard deviations of the experimental and control groups on the two dependent measures are shown in Table 3. The experimental group means were greater than control group means for both the teacher-constructed test and number of squares colored.

Inspection of the test of significance further reveals that interaction of two blocking variables— sex and age— was highly significant. Separate univariate F-tests on each of the dependent variables also yield F-ratios significant beyond the .01 level. The means for each sex and age group are presented in Table 4. The table indicates that the younger boys performed better than the older boys and better than either group of girls. None of the other pairwise comparisons was significant.

STEPHEN BULL, PRIMARY UNIT

During the year 1966-67 an R & I Unit was organized at the Stephen Bull School with the purpose of non-grading the primary grades. Five classrooms with a total of 125 students— six, seven, and eight year olds— were organized into a non-graded primary. Miss Patricia Hansen was the Unit leader. Mr. Frank Sweet was the principal. Other staff members included: Lois Borg, Florence Burcibe, Rosemary McCormack, Jessie Rendall, and Stacy Peterson.

The teachers in this Unit developed an "extended day" plan in the reading program. The youngest and least advanced readers started their day first and went home early while the older youngsters and more able readers came to school later and remained later. The teachers felt this gave them much more time for individual help in reading.

A Mother's Club provided a close relationship between home and school. Evening meetings enabled interested mothers to learn more about a variety of curriculum materials as well as to secure information about the organizational plan of instruction. The mothers appreciated the opportunity to observe their children in some afternoon meetings. A highlight of this club was a Christmas dinner held in the school gym.

Table 2

Tests of Significance Using Wilks' Lambda Criterion

	Treatment	Sex	Age	Treatment x Sex	Treatment x Age	Sex x Age	Treatment x Sex x Age
F	2.479	1.540	1.888	.968	.260	5.842	2.567
df	2/51	2/51	2/51	2/51	2/51	2/51	2/51
Prob.	.094	.224	.162	.387	.772	.005	.087

Table 3

Means and Standard Deviations for the Two Treatment Groups

		Means		Standard Deviation	
Treatment	Number	Squares	Concepts	Squares	Concepts
Experimental	21	99.77	79.11	47.88	32.03
Control	39	72.93	61.69	50.60	32.73
Both Groups	60	81.99	67.52	50.52	33.04

Table 4

Means for Each Sex and Age Group

Age	Test	Boys	Girls
Younger	Squares	131.2	64.1
	Concepts	91.8	56.3
Older	Squares	60.2	85.7
	Concepts	56.0	71.8

Subjects

The subjects for the experiment included the 51 pupils of the non-graded primary Unit. Because approximately half of the students were first year primary students and the remainder second year students, two experiments were actually conducted.

Treatments

The control groups used the Silver Burdette Modern Arithmetic Through Discovery text for Grade 1 or 2 as appropriate, plus teacher-made materials. The experimental group used the same texts, teacher-made materials and the locally-produced skills folder as well as individual conferences to provide knowledge of progress.

Design and Procedures

Students were initially separated into two groups reflecting what normally would be considered grade level. Each group was then stratified according to sex and age—older or younger— prior to random assignment to the experimental or control treatment. Thus, the experimental and control groups at both levels each had approximately a dozen pupils.

Both groups of experimental subjects were combined for instructional purposes; likewise, the control students from both levels met together for arithmetic.

The learning specialist did not teach either group, but worked with teachers in identifying work to be done, in establishing the groupings, in selecting materials, and in providing transition instruction during the teacher rotation between groups. Teachers rotated once half-way through the experiment. The duration of the experiment was ten weeks.

Data Gathered

Different tests were constructed by the teachers for the two levels of students. The

test consisted of a random sample of three items from each concept area in the skill folder.

The tests were administered at both the beginning and the end of the experiment.

Analysis of Data

Because different tests were used at what would normally be the first- and second-grade levels, analyses were performed separately for the two levels. Furthermore, distinct analyses were performed for each group stratified first by sex and then by age. In each case an analysis of covariance was performed on the posttest total, using the pretest composite score as a covariate.

At the first-grade level there was no significant difference attributable to the treatment difference, nor was either stratifying variable — sex or age — significant. A comparison of pretotal and posttotal means, however, indicates that the gain for the experimental group exceeded that of the control group (Table 5). This table also indicates that the boys scored consistently lower than did the girls. This was true as well in all of the subtests administered at the conclusion of the experiment except for one in which there was no difference between boys and girls. Younger students also consistently performed more poorly than did older students on all subtests.

In May, after the experiment had been concluded, Stanford Achievement Tests, Primary I Battery, were administered to all first-year primary students. The mean grade scores on the arithmetic subtest were 1.9 for the experimental group and 2.0 for the control group. These scores indicate that both groups' achievement was satisfactory.

Similar analyses of the second-grade scores again indicated no significant effect attributable to treatment differences. Sex, as a blocking variable, was again not significant, although boys performed more poorly than did girls. One difference, however, between first- and second-grade analyses was the significant relationship between age and performance at the second-grade level. The results of the analysis of covariance are presented in Table 6.

Table 5

Pretest and Posttest Means for First Year Primary Pupils on Teacher-Constructed Test

Group	Sex	N	Pretest Total	Posttest Total	Gain
Experimental	Male	4	40.75	44.50	3.75
	Female	7	56.14	63.00	6.86
	Weighted mean		50.55	56.27	5.72
Control	Male	3	48.33	45.00	3.33
	Female	9	50.44	54.67	4.23
	Weighted mean		49.90	52.25	2.35

Table 6

Analysis of Covariance on Posttest Total Scores of Second Year Primary Pupils

Source	SS	df	MS	F
Treatment	150.288	1	150.288	.90
Age	526.331	1	526.331	3.15 $p < .10$
Treatment x Age	52.358	1	52.358	.31
Error	2672.214	16	167.013	

Table 7

Pretest and Posttest Means for Second Year Primary Pupils, Teacher-Constructed Test

Group	Age	N	Pretest Mean	Posttest Mean	Gain
Experimental	Older	5	81.4	82.8	1.4
	Younger	5	51.0	54.8	3.8
	Weighted mean		66.2	66.8	2.6
Control	Older	7	65.29	57.29	-8.00
	Younger	4	60.75	66.00	5.25
	Weighted mean		63.64	60.45	-3.15

Inspection of the table of means (Table 7) indicates that the younger children made greater improvement than did the older, although their means were in many cases lower than those of the older children. It is also evident that the mean of the experimental group exceeded that of the control group.

If posttest scores are adjusted for pretest performance, the group means favor the younger children. This is the reverse of the age-adjusted score relationship found for the first-year students.

Stanford Achievement Tests, Primary II Battery, were administered in May to the second-year students. In Table 8 the means in grade-equivalents for performance of second-year students on the arithmetic computation and concepts subtests are presented.

Table 8

Second-Year Primary Students' Mean Performance on Kuhlmann-Anderson Intelligence Test and on Stanford Achievement Tests, Primary II Battery, Arithmetic Subtests

Group	N	Mean IQ*	Arithmetic Computation**	Arithmetic Concepts**
Experimental	10	98	2.8	2.2
Control	11	94	2.3	1.9

*Kuhlmann-Anderson, administered in October 1966.

**Stanford Achievement Tests, Primary II Battery, administered in May 1967.

The table indicates that the experimental group outperformed the control group on both subtests. However, this difference is related to the difference in mean IQ of each group. The relatively poorer performance of both groups on the arithmetic concepts subtest than on the computation subtest may be attributable to weakly developed reading skills. In fact, the mean grade equivalent of these 21 students on the paragraph meaning subtest was 2.1.

STEPHEN BULL, THIRD-GRADE UNIT

The third-grade R & I Unit at Stephen Bull School, Mr. Frank Sweet, Principal, continued for the second year under the leadership of Mrs. Mae Elsdon. The teachers of the Unit included Miss Kathy Zaumeyer, Miss Mattie Boykins, and Mrs. Shirley Roth. There were 76 students in the Unit, mostly from backgrounds classified as educationally disadvantaged.

In addition to the extensive programs in reading and mathematics which are described elsewhere, the Unit under the direction of Mrs. Elsdon and with cooperation from the primary Unit and Miss Hansen, organized a Mother's Club with hope of creating a better rapport between home and school. The staff of these two Units in their joint meetings outlined these specific goals: (1) to build a better rapport between home and school, (2) to increase the cooperation of parents, (3) to familiarize the parents with the activities of the school, and (4) to improve the self-image of the mothers by encouraging them to work within the framework of the organization. Plans for accomplishing these goals included personal contacts with the parents through home visits

and phone calls and regularly planned meetings at the school to share ideas and to socialize. The general feeling of the mothers toward this club was one of pleasure at being together. They especially felt the opportunities for learning more about school were especially helpful.

The teachers of this Unit felt that the year's activities had improved the students' attitudes toward school greatly. There were only three cases of absence for reasons other than sickness. For the most part, the children seemed happy in school and were anxious to attend. The teachers observed that the children sought out the Unit leader for help and counseling. Conferences were held regularly before and after school and during school hours when the need arose. The children's problems included requests for help in a subject matter, difficulties with family or peers, and health problems. There was indication of marked improvement because of this personal interest in children's attitudes.

Subjects

The subjects for the experiment were 49 third graders, all eight-year-olds with an average age of 8 years, 6 months. Subjects were stratified by age and sex prior to assignment to experimental and control groups. Table 9 explains the design.

Table 9

Distribution of Children in Experimental and Control Groups

Sex	Age	Experimental	Control
M	Younger	5	6
	Older	6	5
F	Younger	7	6
	Older	6	7
Total		24	24

Treatments

The control group used the Greater Cleveland Mathematics Program (Science Research Associates) with accompanying workbooks and the Scott-Foresman basic text. The experimental group used the preceding commercial material and the Racine-developed material and procedures.

Both the experimental and control groups were given a pretest of items related to each concept to be taught. After the pretest the control group continued to receive conventional mathematics instruction.

The experimental group of children was given an arithmetic folder on which were listed the concepts previously identified. When a concept had been presented, practice work done, and it was evident that the work was generally understood, the children were given a test to assess their understanding of this concept. Each child who received a perfect score on the test could color the box on his folder that indicated the concept tested. He then proceeded to the study of another concept. Those who made errors were grouped for reteaching. Perfect scores were required in order to develop accuracy. The tests were kept as simple as possible.

The experiment covered 40 teaching sessions. At the end of 20 sessions, the teachers rotated. Time allotments were the same for both experimental and control groups. The Unit leader did not teach in any of the sessions. She assumed responsibility in 1) identifying appropriate concepts; 2) establishing groupings of the children based on folders, pretests, and teacher judgments; 3) identifying suitable instructional materials and manipulative devices to provide for the individual differences; and 4) coordinating the teacher rotations, test analysis, and reports.

Analysis of Data

The pretest and posttest mean totals are presented in Table 10 along with the gain scores. It is apparent from the table of means that the experimental group, although initially at an advantage, made greater gains than did the control group. Especially notable was the gain made by the older experimental subjects.

Table 10

Mean Totals for Subgroups on Teacher-Constructed Test

Group	Age	Pretest	Posttest	Gain
Experimental	Younger	83.23	94.90	11.67
	Older	65.08	91.47	26.39
	Mean	74.16	93.19	19.03
Control	Younger	70.09	83.79	13.77
	Older	68.83	84.29	15.46
	Mean	69.44	84.04	14.60

An analysis of covariance was performed on the total posttest scores, using pretest composite scores as a covariate. Table 11 indicates the results of the analysis.

Table 11

Analysis of Covariance on Teacher-Constructed Posttest

Source	df	MS	F
Treatment	1	238.947	2.438
Sex	1	0.527	0.005
Age	1	732.940	6.292 p<.05
Treatment x Sex	1	7.661	0.066
Treatment x Age	1	355.590	3.053 p<.10
Sex x Age	1	10.375	0.089
Treatment x Sex x Age	1	4.996	0.043
Regression	1	21745.474	186.673
Error	34	116.490	

The age difference was significant beyond the .05 level of significance, with older children performing better than younger. However, the relatively large F-ratio for the interaction between treatment and age suggests the appropriateness of looking at means showing the differential effects of treatments for the two age groups. The following table of means adjusted for the pretest scores (Table 12) makes evident both the superior performance of the older children and the large difference in achievement between the younger and older children in the experimental group.

Table 12

Adjusted Means of Subclasses

	Experimental	Control	Mean
Younger	84.08	85.49	84.79
Older	97.98	87.16	92.57
Mean	91.03	86.33	88.68

An a posteriori comparison of means indicates that the older experimental group performed significantly better than any of the other three subgroups. None of the other pairwise comparisons were significant. In other words the older subjects performed significantly better in the experimental condition than they did in the control condition, but there was not a significant difference in performance between younger children in the two treatment groups. Despite adjustment of scores for prior knowledge, older children outperformed younger children. The difference between the two age groups was accentuated in the experimental treatment.

In addition to the pre- and posttesting of achievement an assessment of the children's favorite subjects was made. The children were given this assignment: "If your teacher said that you could have any class that you wanted to start a school day, you would probably choose your favorite. What would be your choice?"

Twenty-two of the total 49 subjects gave mathematics as their first choice.

DISCUSSION OF EXPERIMENTS IN INDIVIDUALIZATION AND MOTIVATION OF ARITHMETIC INSTRUCTION

The three preceding experiments, all conducted in primary classrooms, had as a common element the use of arithmetic folders in which concepts and skills were listed. Pupils using the folders colored in a square when they had mastered a relevant item. It was hypothesized that this procedure, by which a child could evaluate his past accomplishments, would motivate his future performance.

Two of the experimental treatments additionally featured a weekly five-minute individual teacher-pupil conference in conjunction with the use of the folder. In Table 13 a simplified summary of the treatments contrasted in each experiment is found.

In the Howell R & I Unit's experiment, significant differences in favor of the experimental

Table 13

Experimental and Control Treatments in Three Arithmetic Experiments

School	Grade	No folder or conference	Folder only	Folder and conference
Howell	2		Control	Experimental
Stephen Bull	1-2	Control		Experimental
Stephen Bull	3	Control	Experimental	

group were found. The group using both the folder and conference technique performed better than the group not conferring individually with the teacher about the folders.

When the folder-conference procedure was contrasted with neither folder nor conference, no significant difference was found. However, the performance of the experimental groups on the teacher-constructed test was superior to that of the control groups. Since separate analyses were performed for the two grade levels represented in the Stephen Bull primary Unit, the small group size may have contributed to the lack of significance. Given the fact that means were in the expected direction, a larger group size might have made the statistical test adequately sensitive to the differences between groups.

Interesting results came from the folder-no folder comparison in the third grade Unit at Stephen Bull. While the folder procedure was not found significantly better for all children, it was superior for the older children.

Results of the statistical analyses, then, encourage the use of individual evaluation through both the folder and conference technique. The differences in means between experimental groups and control groups favored the former in each of the three experiments, sometimes significantly. Further research is needed to determine which procedure is most appropriate for children of various grades.

Nonstatistical outcomes of the studies must be included in the overall assessment. Of great benefit was the conceptualization by the teachers of a year's arithmetic program. The teachers involved in these studies were generally pleased with the program. They especially saw great potential in the use of an individual record folder. The children were self-motivated to do well in their folders. Also the folder was very useful in reporting to parents.

The idea of having a complete conceptual scheme for the year's mathematics program helped the teachers plan a more unified program. Teachers seemed to be more aware of the importance of some sequential building of skills. Considerable effort was given to developing additional materials.

All three of the Unit leaders were enthusiastic about this project. They commented that they were better informed about the mathematics program and more knowledgeable of the needs of the children. This activity seemed to prove that although we teach material and believe that it is learned by the children, this is often not true. As teachers kept class charts on the work covered, they seemed to plan their work more carefully because they know exactly what was needed by each child. Enthusiasm on the part of children and parents for the reporting techniques was also expressed.

III

INDIVIDUALIZATION IN HANDWRITING AND SPELLING, GIESE SCHOOL

Doris M. Cook, James Wardrop, Mary Quilling, Barbara Kennedy, and Wayne Otto, R & D Center
Maxine Valta and Marilyn Klonecke, Learning Specialists

THIRD-GRADE UNIT

A unique organizational pattern along with a large learning center adjacent to the individual classrooms encouraged a complete individualized program in the third-grade R & I Unit, Giese School, Racine, Wisconsin. The staff of the Unit was composed of Maxine Valta, Unit leader; Janet Hanson and Beverly Schinderle, certified teachers; and Mrs. Louise Weber and Mrs. Lee Ann Paulson, teacher aides. Mr. Earl Nelson, principal, gave constant guidance and support as they tried out these unique arrangements.

Individual word recognition tests were given to all 75 students in September. Each student's reading level was assessed and an appropriate instructional program planned. Miss Mildred Brady, Reading Consultant for Racine Unified District, worked closely with the teachers and made recommendations for both remedial and enrichment work.

Regrouping was done twice a week in mathematics depending upon the achievement level of each child for acquiring the concept being taught. A mathematics corner was organized and set up in the large learning center for independent study. Students who had acquired the basic concepts were encouraged to pursue problem-solving games independently. Other students needing additional practice with certain skills were provided opportunities to manipulate concrete materials and were tutored by one of the teacher aides who supervised the mathematics learning corner.

The teacher aides provided assistance of a clerical and also semi-instructional nature. The reduction of time-consuming clerical tasks gave the teachers more time for better lesson preparation. The teacher aides assisted the teachers in the individualization of instruction and individual record keeping. The teachers felt the instructional program was more effective because of time for planning and teaching.

The two research projects undertaken by the Unit were a developmental study of individualization of spelling instruction and a controlled experiment on various approaches to individualized handwriting instruction, descriptions of which follow.

The teachers of the third-grade Unit were concerned with methods of individualizing their spelling program. The traditional program seemed inadequate. The high ability students could usually spell the words of the weekly unit on Monday without the practice lessons, but the low ability students were having great difficulty with the weekly list. With this information and without any formal testing, the teachers proceeded to develop an individualized spelling program.

Spelling was scheduled into a language arts block of time. The children's stories as well as their written work from other curricular areas became the source of spelling words. Misspelled words were typed onto prescription cards with corrected errors for each individual child. Good spellers were challenged with several new words on their list each week. These words were often taken from other subject areas. For example, many of the students' spelling words included some of the mathematics vocabulary, such as equivalent, equation, etc. Most of these high achievers could study their words independently. The low achievers received additional help from the teachers and the teacher aides.

The basic spelling text, Spell Correctly (Benthal, H. F., Silver Burdett Co.) was used

for a pretest on each Monday. Misspelled words were recorded on each child's individual spelling card. The children who had misspelled words then followed the unit lesson as outlined in the spelling book. On Wednesday the children tested each other on these words. Those who had mastered these words could then have additional words.

The instructional program incorporated teaching techniques appropriate for both low achievers and high achievers. Thus the individual needs of all students were met.

While no pretesting was done specifically to assess these results of the program, the teachers indicated many observations of the effectiveness. They cited (1) improved interest, (2) the motivational effects of concrete rewards, (3) interest shown by parents helping students at home, (4) enthusiasm of the children for spelling and specifically for having their own spelling card, (5) improved ability of the low ability students to sound out words, (6) the usefulness of the teacher aides helping low achievers.

THE COMPARATIVE EFFECTS OF THREE METHODS OF HANDWRITING INSTRUCTION UPON PUPIL SKILL IN HANDWRITING

Experimentation with handwriting instruction was begun at Olson School, Racine, Wisconsin, during the second semester of 1965-66 (see Hieronymus et al., 1967).[1] Miss Elizabeth Williams, Language Arts Consultant, Racine Unified District, provided valuable consultant help during the planning and experimentation. Dr. Glenn Dupuis, postdoctoral fellow, R & D Center, assisted in the experimental design for the study. The early experiments compared the effect of various instructional approaches upon the development of pupils' handwriting as a communication tool.

Problem

The 1966-67 experiment continued with the primary objectives of creating a desire in the students to do their best handwriting throughout the day. Results from the early experiments as well as teachers' reactions toward the individualized approach, it was decided to explore various ways of implementing an individualized program, with emphasis on self-evaluation and individual progress. The specific purpose in this experiment was to determine the relative effect of three handwriting treatments—(1) the traditional approach, (2) an individualized-instructional approach, and (3) a comprehensive individualized approach—upon selected measures of pupils' handwriting performance.

Subjects

The subjects were 72 third-grade children at Olson School. The experiment began late in September when the students seemed ready for the transition from manuscript to cursive writing.

Experimental Treatments

The three treatments utilized different materials as well as different methods of instruction, as can be noted in the descriptions which follow:

1. Traditional Approach. This treatment is the traditional approached outlined by the Racine Unified District. Adventures in Handwriting, Peterson Directed Handwriting, published by the Macmillan Company, was used as a basic text. All pupils in this treatment practiced the same skills as outlined by the district's Handbook of Handwriting Practices. The students followed teacher-made models from the chalkboard. All small and capital letters were reviewed and practiced in words and sentences.

2. Individualized Instructional Approach. For this treatment no commercially prepared program was used. The Peterson letter formations were used as models, but the practice exercises were prescribed by the teacher and written on each student's individual practice record. The techniques used for practicing handwriting were determined on the basis of teacher judgment of individual needs. Much of the handwriting practice was coordinated with composition and spelling. After ten minutes of practice the children were given a creative writing assignment which was then used by the teacher as an evaluation device for the following lesson. Samples of written work from all subject areas were constantly evaluated by both teacher and student. Effort was made to influence the students to do their best handwriting throughout the day.

3. Comprehensive Individualized Approach. This treatment differed from treatment 2 in several ways:

a. Two teacher aides assisted the Unit leader in individual and small-group practices.

---

[1] Hieronymus, A. N., Cook, Doris W., W. L., Dupuis, G. E., and Fingal, L. Individualizing instruction in language arts through development and research in R & D Units of local schools, 1965-1966. Technical Report from the Wisconsin R & D Center for Cognitive Learning, University of Wisconsin, 1967, No. 12.

b. The overhead projector was used to show models of letter formations and also to present lessons.
c. A commercially prepared programed course, The Penskill Individualized Handwriting Skills Program, published by Science Research Associates, Inc., was used as supplementary material for a group of high-performance children.
d. In an effort to encourage self-evaluation and individual progress, a folder was kept for each child. Samples of the child's writings were kept in this folder along with a practice record schedule. The staff, when conferring individually with the students, indicated areas that needed practice: letter forms, size, space, and slant. The child then proceeded independently, and as improvement was shown this was recorded in his folder by indicating the date. In addition to being a practice record, this folder was used for periodic recording of the student's progress on his writing the sentence "The quick brown fox. . . ." Each sample was dated, and the child had an opportunity to see his improvement.

Procedures and Design

Students were stratified by sex and then randomly assigned to one of the three treatments. Table 14 indicates the make-up of the three groups at the end of the experiment. The lack of porportionality of males to females reflects the fact that complete data were not available for all students initially assigned to the three groups. The Unit leader had responsibility for presentation of each instructional program. She taught one handwriting class each day. In other words, every third day a student had a 30-minute formal handwriting class. For treatment 3 the instructor was assisted by two teacher aides. The experiment ran for 27 weeks.

Table 14

Distribution of Children in Experimental and Control Groups, Giese School, Grade 3

Treatment	Sex	
	M	F
Traditional	11	13
Individualized	9	14
Comprehensive Individualized	14	10

Data Gathered

Handwriting samples were collected from each child under three conditions both before the treatments commenced and at the conclusion of the experiment. The three conditions are described below:

1. Normal Condition: Teachers said to their students: "Write the sentence 'The quick brown fox jumps over the lazy dog' as you usually write. Keep writing the sentence until I tell you to stop. I will pass out a sheet of paper with the sentence printed on it so that you can spell all the words correctly. Now write your name on your paper, but do not start writing the sentence until I tell you to begin. Leave a blank line between sentences." Once each child had a slip of paper with the sentence printed on it and pencil and paper with his name on it, he was told to start. At the end of five minutes he was told to stop. The papers were collected and filed in the "normal" folder. The students kept the piece of paper with the sentence printed on it.

2. Best Condition: The teachers said to their students: "Now write the same sentence as well as you can. Write the best you can. Write your name at the top of the next sheet of paper and then write the sentence four times as well as you can. Leave a blank line between sentences."

3. Fastest Condition: The teachers said to their pupils: "Now write the sentence as fast as you can. Keep writing the sentence until I tell you to stop. Now write your name on your paper, but do not start writing the sentence until I tell you to begin. Remember to keep writing the sentence over and over as fast as you can. Leave a blank line between sentences." As soon as the children were ready, they were told to start. At the end of five minutes they were told to stop. The papers were then collected and put in the "fastest" folder.

A handwriting sample from another curricular area was also obtained.

Scoring and Analysis of the Data

Three judges rated the pre- and posttreatment samples for legibility by using a 7-point scale devised for another study[7] and described in

[7] Otto, W. and Erwick, C. L. Effect of time of transition from manuscript to cursive writing upon subsequent performance in handwriting, spelling, and reading. Technical Report from the Wisconsin R & D Center for Cognitive Learning, University of Wisconsin, 1968, No. 67.

detail in that report. The naturalistic samples from another curriculum area were rated by the same judges, but a formal scale was not used; instead the samples were simply sorted into seven legibility categories. The validity of this approach is discussed elsewhere.[3] The legibility ratings assigned by each judge were totaled and averaged for each sample, and the resulting mean scores were used in the analyses. Total words produced by each subject under the "normal" and "fastest" conditions were also used in the analyses.

Data were analysed using a two-way analysis of covariance, with two covariates for all measures except "Sample from other Area." IQ was used as a covariate for all measures, with the corresponding pretest measure as the second covariate where appropriate.

A summary of the analyses is given in Table 15. The only significant treatment effects were for the measures of speed of handwriting under "normal" and "fastest" instructions. Table 16 gives the means and standard deviations for this factor, and indicates that the Traditional group was the slowest in both cases, while the Individualised and Comprehensive groups were significantly faster under "fastest" instructions.

[3] Otto, W., Askov, Eunice, and Cooper, Carin. Legibility ratings for handwriting samples: a pragmatic approach. Perceptual and Motor Skills, in press.

Table 16

Means and Standard Deviations for Significant "Treatment" Effects, Grade 3

Treatment	Dependent Variable			
	No. of words Normal		No. of words Fast	
	Mean	SD	Mean	SD
Traditional	31.54	14.18	48.54	14.60
Individualised	35.09	11.05	52.48	13.93
Comprehensive	43.75	14.42	51.04	13.04

The only other significant effect was a Treatment x Sex interaction for "Number of words fastest." This interaction seems to result from the marked superiority of the females in the "Individualised" treatment. (Unadjusted means are presented in Table 17).

The means for each group under the normal condition are also of interest, because they may be compared with means gathered in a statewide sample in which the same scale was used to judge handwriting.[4] Furthermore, two of the three judges used were common to the two studies. The comparative statewide data

[4] Otto, Askov, Cooper, op cit.

Table 15

Summary of Analyses of Covariance, Grade 3

Source of Variation	df		Dependent Variable					
			No. of words normal	No. of words fastest	Normal rating	Best rating	Fastest rating	Sample from other areas
Treatment (T)	2	MS	711.08	488.12	0.57	0.37	0.72	0.65
		F	5.18**	3.87*	<1	1.19	1.06	<1
Sex (S)	1	MS	92.31	15.87	0.19	0.78	0.12	5.52
		F	<1	<1	<1	<1	<1	2.78
T x S	2	MS	281.45	494.79	0.81	0.14	0.88	1.04
		F	2.11	3.87*	1.35	<1	<1	<1
Regression (Covariates)	2	MS	955.32	2192.36	9.19	10.32	5.12	3.81[1]
		F	6.91**	15.58***	15.82***	13.51***	3.69*	1.95[2]
Error	60	MS	137.92	140.78	0.57	0.81	1.38	2.18[2]

* $p < .05$
** $p < .01$
*** $p < .001$

[1] df = 1
[2] df = 60

Table 17

Unadjusted Means for Treatment x Sex Interaction, Number of Words Fastest Measure

	Sex			
	Males		Females	
Treatment	Mean	SD	Mean	SD
Traditional	39.91	17.07	41.08	10.78
Individualized	39.33	14.92	40.93	12.16
Comprehensive	48.07	15.06	55.20	7.54

were gathered in February of fourth grade, at which time students would presumably be writing better than would third graders. The means for the two studies are presented in Table 18.

Comparison of means shows that both boys and girls performed better at Giese School than those children rated in the statewide sample. The only group at Giese which did not outperform all of the statewide groups of the same sex was the group of girls receiving the traditional treatment. Their average rating differed by less than 0.10 from other girls who had made the transition to cursive in third grade.

Discussion

While significant effects were noted in favor of the experimental groups only on analyses of speed data, the mean ratings under normal conditions of experimental groups also exceeded those of traditional groups. Furthermore all groups performed as well as or better than a statewide sample of fourth graders writing under the normal condition. This latter result suggests that the experiment contributed to excellence of handwriting in all groups in Giese School: the traditional treatment differed from the instruction many third graders receive because of the instructor's awareness of experimental treatments. Thus, the traditional approach may have been taught with an excellence that mitigated the difference between means of this group and either experimental group.

Also of interest is the consistent sex difference both in speed and quality of handwriting. The difference favors the girls in all cases—a finding consistent with earlier research.

Implications

Of special consideration are the staff's opinions and reactions to this year's program. The Unit leader and teachers felt that an individualized method might best be implemented after basic letter formations are introduced and taught. This could be done in group instruction following the plans of the chosen manual for letter formation. In this case it was the Peterson Adventures in Handwriting.

After all the basic letter formations have been taught, then an individualized approach could proceed. The individual folders could be introduced at this time and could become the basis for independent study. The staff felt the individual records proved very worthwhile for several reasons:

1. The children responded enthusiastically to having their own progress reports.

2. The children could pursue their writing practice with more independence. This eliminated much teacher direction.

Table 18

Normal Condition Ratings for Giese Third Graders and a Statewide Sample of Fourth Graders

	Group "Treatment"		Girls	Boys	Mean
Giese	Traditional		3.819	3.606	3.713
	Individualized		4.024	3.777	3.901
	Comprehensive Individualized		4.267	3.691	3.979
	Mean		4.037	3.691	3.864
Statewide Sample	Transition to Cursive	1st Semester of 2nd. Year	3.533	1.777	2.655
	Transition to Cursive	2nd Semester of 2nd. Year	3.777	2.357	3.067
	Transition to Cursive	1st Semester of 3rd. Year	3.666	2.667	3.166
	Transition to Cursive	2nd Semester of 3rd. Year	3.800	2.733	3.312
	Mean		3.767	2.384	3.073

3. The children responded positively to receiving constant feedback about their writing.

4. The folders encouraged self-evaluation.

5. The teachers had a quick and concise account about each student's needs. Appropriate instruction could then be provided.

Papers from other areas of the curriculum could be collected regularly and added to the folder to be used by the child for his practice. Having these folders kept in a specific place in the classroom available at any time eliminated the need for a formal handwriting lesson every day. However, if a teacher felt a certain letter combination was being written incorrectly by a majority of the class, he could include a short drill or practice in any of the language arts. This was often done with spelling in this study. Experiences indicate that students having extreme difficulty benefited greatly from the individual folder where they could see their difficulties and what they needed to practice.

After the basic formations had been taught, students of high performance could proceed on their own in the Penskills, a commercially individualized program.

The teachers felt they assumed quite a different role in the individualized approach. The time spent usually on presenting letter formations to all students was used much more effectively. Large group presentation with the use of audio-visual materials gave the teachers time to circulate among the students and give help where needed as well as assist in their independent activities. One very important role of the teacher was going over each individual folder and circling poor letter formations with a colored pen and writing the correct form on the child's record where it served as a model for his practice writings. This is very time consuming and probably is impossible without the additional help provided by the use of teacher aides who can provide assistance to individual students' needs as outlined by the teacher.

FOURTH-GRADE UNIT

The fourth level R & I Unit at Giese School in Racine attempted to combine the advantages of a physical facility constructed for a team approach to instruction with the benefits of cooperative planning and teaching under the guidance of Mrs. Marilyn Kletecka, Unit leader. Other staff members were Mr. Charles Leonard, Miss Anne Buchanan, and Miss Mary Rounds, certified teachers.

After assessing the achievement of the children in the Unit, the classroom (homeroom) groups were reorganized for instruction in reading, spelling, and arithmetic. By employing the learning specialist as a teacher for one group and by using the large group instruction area as a teaching station, four instructional groups could be formed for combined reading-spelling classes. The formation of two rather large groups of about 30 high and 30 average achievers allowed for the formation of two much smaller groups of 15 slower readers. The teachers then had the responsibility of planning only for one reading group; however, since within these relatively homogeneous groups there still were many individual differences, some of the teachers formed two groups, and all teachers used the SRA Reading Labs. A basic spelling text was used for all groups; however, the high achievers moved quickly through the basic text list words and included the week's challenge words suggested by the text, plus words from their readers which interested them, whereas the low achievers often took only part of the text list for a week's work.

Compared with reading, differences in mathematics achievement were equally pronounced. Again with the Unit leader as a teacher for one group and with the large group area for instruction, the Unit was able to individualize mathematics. About 40 students were identified who (1) had a good command of the basic number facts and of addition-subtraction computation skills, (2) had a reading ability of approximately 3.0, and (3) displayed the maturity necessary for self-directed learning; these students entered IMCP—the Individualized Mathematics Curriculum Project planned by Dr. M. Vere DeVault of the University of Wisconsin. In this program the children chose their own area of interest in mathematics (such as sets, geometry, multiplication-division), checked and recorded the results of their work, and progressed at their individual rate. Two teachers supervised this program—answered questions, further explained directions when necessary, administered the tests which are a part of the program sequence, and, most important, held individual conferences with students to discuss and guide their progress.

Another group of about 30 students who were achieving in mathematics at approximately grade level but who did not have the reading skills or the maturity necessary for IMCP were taught by one teacher who used the basic text series and followed a "normal" fourth level sequence. The third group of students was comprised of low achievers; efforts were made to make mathematics as enjoyable as possible for these children by providing many arithmetic games, concrete visual aides, and "sure-

success" assignments. Emphasis was placed on increasing their knowledge of basic facts and simple computational skills.

Several times during the year when the Unit teachers wished to introduce a particular mathematical concept to all the Unit children at the same time, the classes were taught as homeroom groups. Once a week homeroom groups met to work specifically on basic facts and computational skills; drill in these areas was carried out daily. The Unit teachers wished to be aware of progress in all three mathematics groups so they rotated teaching responsibilities every 3 weeks. And, whenever it seemed to a child's advantage, he was moved from one group to another.

Responsibility for planning and teaching social studies and science units was divided among the teachers. Usually one or two teachers planned the entire unit of study and the others offered suggestions for improvement. All units were arranged with an initial large-group presentation by one teacher as motivation for the following small-group activities. During these presentations, as well as during films and filmstrips throughout the unit of study, at least two teachers were free to prepare materials for the follow-up activities or to pursue further planning. The range of student abilities was stressed in planning these small-group activities; there were many nonreading projects suggested for the less able students, while more able students could choose more difficult independent activities. In preparing their projects, which included charts, graphs, murals, models, reports, maps, plays, experiments, demonstrations, tape recordings, and illustrations, students had free access to the school resource library as well as to the books, pamphlets, pictures, objects, and audiovisual equipment provided in the learning center (in the large-group instruction area). The list of suggested activities was always extensive enough to allow the student to move immediately to a new activity after completing another.

During the second semester the Unit teachers had an opportunity to teach science classes using the AAAS approach. The processes used in science rather than scientific facts were emphasized. Lessons which were learning experiences for the Unit teachers as well as for their students included making and revising predictions, and distinguishing between fact and inference. The children, who adapted eagerly and easily to this approach to science, found the opportunity to manipulate materials and to graph their individual results particularly stimulating. This R & I Unit is looking forward to extending this program throughout the coming year. It is hoped that the entire staff of Giese School will work toward developing an exemplary instructional program in science.

Since classroom teachers in Racine are wholly responsible for instruction in art, music, and physical education, the Unit teachers felt a better job could be done in each of these areas if a single teacher prepared a lesson for only one area at a time and taught the same lesson on successive days in the three homerooms. A pattern was established in which the general topics and sequence of instruction were planned by the entire Unit at its weekly meeting; the individual teachers then planned specific lessons. For example, when the Unit decided to emphasize printmaking in its next art classes, one teacher agreed to plan for and teach sandpaper-printing, the second volunteered to work with potato-printing, and the third took clay-printing. At times the Unit planned large group activities in these three areas, such as "sings," art films and follow-up art activities, and folk dancing.

In evaluating the teaching of language arts, the Unit felt its program was somewhat haphazard. Subsequently, an attempt was made to organize the language arts program by focusing on one aspect of language instruction each day of the week. The areas established were listening, speaking, composition, grammar, and library skills. This plan seemed to give more direction to the Unit's language arts program than simply following the basic text had done. Aspects of the language arts program such as spelling and composition were incorporated into the handwriting experiment described on the next pages.

THE COMPARATIVE EFFECTS OF THREE METHODS OF HANDWRITING INSTRUCTION

The specific purpose in this experiment was to determine the relative effect of three handwriting treatments— (1) the traditional approach, (2) the commercial individualized approach, and (3) the noncommercial individualized approach— upon selected measures of pupils' handwriting performance.

Subjects

The subjects were 78 fourth-grade children at Giese School. Pupils were randomly assigned to one of the three treatments following stratification by sex and by three levels of achievement. The achievement levels were determined from IQ scores and scores on the Stanford Achievement Test administered in September, 1966.

Experimental Treatments

(1) Traditional approach. Commercial materials, The Correlated Handwriting Series, Frank N. Freeman, author, published by Zaner-Bloser Company, were utilized.

(2) For the commercial individualized approach The Penskill Individualized Handwriting Skills Program, by R. G. Larson, published by Science Research Associates, Inc., was used.

(3) No commercially prepared material was used for the noncommercial individualized approach. The techniques utilized with each student were determined on the basis of teachers' judgments of individual needs. Upon mutual agreement among the teachers this general procedure followed: (1) Each child was supplied with an individual folder which was used for self-evaluation of his writing materials. Three-eighth-inch ruled paper was used by all the students for both practice and testing. The students' writings in other areas of the curriculum were the basis for the handwriting instruction. (2) The correctness of the letter formation was based on legibility and did not necessarily conform to a letter formation advocated by any commercial producer. Self-evaluation was carried on daily and continuously by the student and the teacher. The opaque projector was used to teach pupils how to evaluate handwriting. (3) Pupils received instruction through indication of individual needs for remedial or practice suggestions. Every three weeks the children were instructed to write the sentence "The quick brown fox jumps over the lazy dog" three times for self-evaluation using the SRA evaluation chart. After deciding their quality level—1, 2, 3, 4, or 5—the students then indicated their score and filed samples in their individual folders.

Procedures and Design

A stratified random sampling plan was used. Students were separated by sex and level of IQ and randomly assigned to one of the three groups. Instruction for each group was provided for 20 minutes three times weekly for 27 weeks. Teachers followed the procedures outlined by the publisher in the commercial programs. Teachers were rotated every four weeks among the three treatments. The design employed in this study is presented in Table 19.

Table 19

Distribution of Children in Experimental and Control Groups, Giese School, Grade 4

Treatment	Ability	Sex M	Sex F
Commercial Individualized	H	N=5	N=4
	M	N=5	N=5
	L	N=4	N=3
Traditional	H	N=4	N=4
	M	N=6	N=6
	L	N=4	N=2
Individualized	H	N=6	N=4
	M	N=5	N=5
	L	N=4	N=2

Data Gathered

Scores were obtained in the manner described in the "Data Gathered" section of the third-grade report. Of the five scores obtained before and after the experiment, two were measures of speed and three were legibility ratings obtained under "normal," "best," and "fastest" instruction. All samples were rated independently by three trained judges, and the mean of the three ratings was used for analysis. Reliability coefficients were calculated for these average ratings for two of the measures, as a check on the accuracy of the judges. For the "normal" rating, pretest, this value was .70, and for the "best" rating, posttest, it was .84.

A sixth measure was taken at the conclusion of the experiment, a sample of the subjects' handwriting from some other curricular area. This sample was rated for legibility in the same manner as the other samples.

Results and Discussion

Table 20 presents a summary of the analyses of variance and covariance which were performed. Analysis of covariance was used for all dependent variables except "Sample from other area." In this case an analysis of variance was employed. For analysis of covariance, the covariate in each case was the corresponding measure from pre-experimental writing tasks.

Significant main effects were found for Ability Level for sample from other area and for Sex for normal rating, fastest rating, and sample from other area. Table 21 gives the means for the significant Sex effects, and indicates that the females were superior in each case.

Means for significant Ability Level effects for sample from other areas are given in Table 22. The high ability group performed

Table 20

Mean Squares and F Ratios for Speed and Legibility of Handwriting, Grade 4

Source of Variation	df		Dependent Variable: No. of Words Normal	No. of Words Fast	Normal Rating	Best Rating	Fastest Rating	Sample From Other Areas
Treatment (A)	2	MS	34.74	18.61	0.37	0.00	0.53	1.92
		F	<1	<1	<1	<1	<1	1.29
Ability level (B)	2	MS	21.72	33.27	2.56	1.12	0.07	9.28
		F	<1	<1	2.08	1.71	<1	6.22**
Sex (C)	1	MS	44.08	26.70	8.78	0.38	5.83	39.48
		F	<1	<1	7.13**	<1	5.05*	26.45***
AB	4	MS	93.80	138.94	0.80	1.86	1.48	1.46
		F	1.62	2.02	<1	2.85	1.28	<1
AC	2	MS	207.78	89.08	1.63	0.73	1.96	1.37
		F	3.60*	1.30	1.33	1.12	1.70	<1
BC	2	MS	25.21	181.51	0.86	1.34	0.47	0.18
		F	<1	2.64	<1	2.05	<1	<1
ABC	4	MS	7.08	30.52	1.08	0.95	0.69	0.86
		F	<1	<1	<1	1.45	<1	<1
Covariate	1	MS	557.00	2045.89	2.66	32.28	23.47	-[a]
		F	9.64**	29.78***	2.16	49.39***	20.30***	
Error	59	MS	59.77	68.69	1.23	0.65	1.16	1.49[b]

* p <.05
** p <.01
*** p <.001

[a] No covariate used for analysis of sample from other area

[b] df = 60

Table 21

Means for Significant "Sex" Effects

Sex	Dependent Variable: Normal Rating	Fastest Rating	Sample from Other Area
M	3.06	2.85	3.07
F	4.20	3.89	4.63

Table 22

Means for Significant "Ability Level" Effects for Sample from Other Area

	Ability Group: H	A	L
Mean	4.51	3.77	3.07

substantially better than either the average or low ability groups.

The significant Treatment x Sex interaction for number of words normal is indicated in Table 23. This interaction seems to result from the marked superiority of the males on the individualized approach and of the females on the traditional approach. The means for each of the treatment groups under the "normal" instructions are also of interest because they may be compared to results from the statewide sample of fourth graders previously described.

Table 23

Cell Means for Significant Treatment x Sex Interaction for Number of Words Normal

	Treatment		
Sex	Commercial Individualized SRA	Traditional Zaner-Bloser	Individualized
M	30.71	29.57	36.00
F	31.83	37.17	31.91

The statewide data were gathered from fourth graders in February and were judged against the same scale by three judges, two of whom served as judges for the Giese study. The means are presented in Table 24.

Discussion

The results of tests of significance indicated that girls' handwriting under several conditions was rated significantly higher than boys, and that rating of handwriting in daily work is directly related to the ability of the child, as measured by an intelligence test. Neither of these results is startling.

Of greater interest is the fact that pupils in Giese School, both boys and girls, performed better, on the average, than did pupils in a statewide sample of fourth graders. Furthermore, both groups whose instruction was individualized received higher ratings under the "normal" condition than did the traditional group. The preceding results indicate that the experiment had a positive effect on students' handwriting, and that individualization was successful, despite the lack of significant differences between treatments.

Table 24

Normal Condition Ratings for Giese Fourth Graders and a Statewide Sample of Fourth Graders

Group	"Treatment"		Girls	Boys	Mean
Giese	Traditional		3.564	2.537	3.050
	Individualized		3.833	2.829	3.331
	Commercial Individualized		4.421	2.622	3.522
	Mean		3.939	2.663	3.301
Statewide Sample	Transition to Cursive	1st Semester of 2nd. Year	3.533	1.777	2.655
	Transition to Cursive	2nd Semester of 2nd. Year	3.777	2.357	3.067
	Transition to Cursive	1st Semester of 3rd. Year	3.866	2.667	3.268
	Transition to Cursive	2nd Semester of 3rd. Year	3.890	2.733	3.312
	Mean		3.767	2.384	3.075

IV

INSTRUCTIONAL PROCEDURES IN LANGUAGE ARTS FOR DISADVANTAGED CHILDREN

Doris M. Cook and James Wardrop, R & D Center
Mae Elsdon, Joe Dahlby, Mary Kilgore,
and Dawn Kloften, Learning Specialists

WRITTEN EXPRESSION, WINSLOW SCHOOL, FIFTH GRADE

The R & I Unit under the leadership of Mr. Joe Dahlby, Unit leader, and David Sweeney, Principal, was in its second year at Winslow School. Mr. John Thomas and Miss Margaret Henze, teachers, assisted in developing an instructional program for 48 students, most of whom were culturally disadvantaged. The staff met regularly once a week and during lunch hours almost daily to share ideas and plan cooperatively.

The staff experimented with a variety of organizational plans and instructional approaches. In reading, interclass groupings provided an opportunity for a staff member to give additional help to the low achievers. One staff member was also available to supervise enrichment activities for the high achievers and talented readers. Study carrels were installed in the wide halls of this old building to provide individual reading centers. An agreement with the city library gave these students many additional reading materials.

Children identified as below grade level in spelling were given individualized instruction in spelling, with SRA Spelling Laboratories the principal resource. Students were tested monthly in their individual lists, which were constructed from misspelling found in their written work.

The need to improve both the content and the mechanics of student writing was evident to the teachers of this Unit. Typically, schools have attempted to improve content by having students write about the subject matter presented to them and to meet the need for mechanical improvement by assigning exercises in grammar, usage, punctuation, spelling, and capitalization. Whether or not content and mechanics improve from these approaches has been frequently questioned. The awareness of these needs led to developmental activities reported here. The theoretical position from which this study started is that the content and mechanics of student writing can be improved simultaneously by stimulating students to inject creativity into their writing and by motivating them with a specific purpose for writing.

Uncaptioned cartoons drawn by the learning specialist were presented to the students in a sequence that suggested a plot. Students were asked to supply a dialogue for the cartoon characters and then write the story the sequence suggested to them including the dialogue they had supplied. The written products were then evaluated by the teachers who suggested content and mechanical revisions, wrote a positively reinforcing comment on each paper, and returned the papers to the students for revision.

Motivating the revision of papers was a source of concern to the Unit staff. Accordingly, a pilot study was planned to evaluate the relative efficacy of two motivational techniques. The students in each of the fifth grade classes involved were randomly assigned either to have their papers typed and filed in their personal folders or to have their final copies displayed on the bulletin board. The Stanford Achievement Test given in October provided some baseline data about the students' ability with the mechanics of writing: spelling, usage, punctuation, capitalization, sentence sense, and dictionary skills. A second form of this test was given in April. The pre- and posttest scores were analyzed by an analysis of covariance to determine whether significant differences between the scores of those students who had

their writing typed for their personal folders and those who had their writing displayed on the bulletin board existed.

While the results of the analysis suggested there was no difference in the relative effectiveness of the two procedures, the response of teachers and pupils alike to the language program was enthusiastic.

These reactions on the part of the students were noticed by the staff. (1) Willingness to write. No child ever failed to do the lesson when he was present. (2) Quantity of writing. The average output per student was a page to a page and a half for a half-hour's writing time. (3) Subject preference. Writing was placed first in an index of subject preference when totaling choices made by all students in the two fifth-grade classes. (4) Handwriting. The quantity of writing and rewriting produced gains in neatness and legibility for most poor handwriters. (5) Creativity. Many children previously considered relatively unimaginative, bloomed as writers during the year.

Not surprisingly, achievement test subscores from fall and spring administrations of the Stanford Achievement test indicate above average gains on those subtests most directly related to this program—spelling and word meaning. The table of mean grade equivalents before and after this language program (Table 25) show gains ranging from seven months to over one year.

Table 25

Stanford Achievement Test Mean Grade Equivalents Fall and Spring, Winslow School, Grade 5

	Fall	Spring	Gain
Word Meaning	5.1	6.2	1.1
Spelling	5.4	6.3	0.9
Language	5.3	6.0	0.7

The table further indicates that the class means were considerably above the national average (5.0) for spring administration. Taking into account the fact that this student population was considered disadvantaged, such growth confirms the effectiveness of the language program. More objective data on improvement in content and creativity available, even stronger evidence for the merits of the program might be presented.

THE EFFECTS OF MOTIVATIONAL PROCEDURES ON CHILDREN'S READING, STEPHEN BULL, THIRD GRADE

The lack of pupil interest in independent reading in the third-grade R & I Unit at Stephen Bull School, Racine, led to the development of an experiment to test the following: Will a system of rewards combined with individual reporting increase independent reading by children? Will this in turn improve reading achievement? Learning Specialist Mae Ekdson, teachers Mattie Boykins, Anne Dodge, and Shirley Roth, assisted by social volunteers from local women's organizations participated in the project, which was carried on from September, 1966, through May, 1967.

Subjects

Subjects for this project were the 72 third-grade children at Stephen Bull School, Racine, Wisconsin. They ranged in age from 7.10 to 10.5 years, the average being 8.6 years. The mean IQ was 97.4. This school is in the inner city disadvantaged area.

Treatment

This program of reading instruction was added to regularly scheduled reading classes. The main objectives were to stimulate an interest in independent reading and to get children to read more. To this end, children had to be provided with books they could read. At the start of the project, many of these children were reading at the preprimer and primer level, and several were nonreaders.

Three motivational techniques were utilized in this program. A brief description follows:

(1) Materials and books appropriate for the various reading levels were made available in the classrooms and the school library.

(2) Concrete rewards were used systematically. A reward system was set up whereby, after reading two books and reporting on them, the child was given a new pencil; after five books, an eraser; after ten books, a box of crayons; after fifteen books two Tiny Golden Books, and after twenty books, a ball point pen. After thirty books had been read, prizes were awarded in multiples of ten.

(3) Individual conferences were conducted by volunteer aides. One afternoon per week volunteer aides visited each classroom to discuss with each child the books he had read and to listen to him read. A child could also gain credit for the book report by writing a letter to the Learning Specialist describing the

book read, or by drawing a picture to illustrate the book. The Learning Specialist also held individual conferences for book reports during her free time. Each child received from 5 to 15 minutes of individual attention when making a report. Despite suggestions for varying the reporting procedure, the children preferred individual conferences and oral reading.

A list of books read by each child was kept by the volunteers. Books read were then recorded on an individual chart and graph where each child could watch his achievements.

Data Gathered

The data gathered for this project included relevant scores from the Stanford Achievement Test, given in March; the Gates-MacGinitie Reading Test, given in February and again in April; teacher-development inventories about pupils' attitudes; and a count of the number of books read.

Results

Since this reading program was not designed as an experiment, but was basically an innovative developmental activity, no formal analyses of the data were conducted. Instead, much of the evaluation of the project was based on teacher judgments and case histories of individual students. Some objective test information was collected, and is reported below.

The mean grade equivalent scores for the class on the Gates-MacGinitie Reading Tests, Primary C, in February were 2.6 for the Vocabulary and 2.7 for Comprehension. In April, the mean for both scores was 3.3. Thus, in two months these students had gained the equivalent of 8 months in Vocabulary and 9 in Comprehension. These gains become even more meaningful because of the fact that these students were consistently below grade level on both measures to begin with, so that it would have been predicted that they would have gained less than 2 months in this time interval.

The Stanford Achievement Test was given in March, as a part of Boston's city-wide testing program. On the Word Meaning subtest, the mean grade equivalent score was 3.1, while on the Paragraph Meaning subtest the mean was 2.8. These values represent gains of 1.3 years on the Word Meaning subtest and 1.2 years on the Paragraph Meaning subtest over scores obtained at the beginning of second grade (a time interval of 7 months). This is a very encouraging result, because children such as these typically fall further and further behind each year. They were already 1.8 years behind in Word Meaning and 1.2 years behind in Paragraph Meaning at the start of grade two, so in this group the trend has at least been halted for a time.

The 72 students reported on a total of 2078 books, the median being 23. Nine children read over 50 books each, with three reporting on 70 or more. Only one child gave no book report.

One child who was a nonreader at the beginning of the year was enthusiastic about reading in the spring. He gained over 20 months in Word Meaning on the Stanford Achievement Test. His comment during a report period one day was, "I'm gettin' better, ain't I?" Other children in the unit also made great improvements (one girl gained 2.9 years in Word Meaning and 3.2 years in Paragraph Meaning on the Stanford), but most of the changes have been of lesser degree than these.

Teachers reported that the children were eager to come to school, and even asked to be allowed to come early or stay late. When an assignment was given in which children were asked, "What is happiness to you in school?" several answers were given like these:

"Happiness is reading."
"Happiness is book reports."
"Happiness is when Mrs. _____ hears me read."

Conclusion

The project described here was an attempt to develop an interest in reading and a desire to read in children of the inner city. Although no formal evaluation was made, the data which were available indicated that the effort was quite successful. Children read many books, responded enthusiastically to the program, and showed a generally improved attitude towards school.

One might prefer that these statements be supported by quantitative experimental data, but innovation and the improvement of the educational programs in the schools arise in many ways. If several students benefit from a program like this one, it is worthwhile. That virtually an entire group has benefitted is all that should be asked.

The concrete rewards which were given may have been sufficiently motivating for some students, but the opportunity to converse with an adult was probably much more meaningful. The opportunities for future research to determine the importance of the various motivational techniques employed are, of course, many.

IMPROVING ORAL AND WRITTEN EXPRESSION, FRANKLIN SCHOOL, SECOND GRADE

The K & I Unit in the Franklin School, Mr. John Stickle, principal, was reorganized in the fall of 1967 under the leadership of Miss Mary Kilgore, Unit leader. She was assisted by Mrs. Ellen Ilse, Mrs. Peter Van Dieser, Mrs. Jean Laufman and Miss Kathy Carolina, teachers.

There were seventy-one students in the Unit, most of whom were classified as educationally disadvantaged.

The students were grouped in reading based on informal inventories and subtests of the Stanford Achievement. There was a range from the readiness (pretesting) to 3.2 grade level. Twelve students who were nonreaders were put into an IPA reading group until February, after which the change to traditional orthography, or the usual alphabet and spelling, was made. All completed a primer and read part of a second semester of first-grade reader. Three of these "slow starters" also read a portion of a second-grade reader.

The low achievers in mathematics were given one additional semester of manipulative math experiences. No textbook was used and no formal instruction of basic skills was given; rather, emphasis was on developing basic understandings. At the beginning of the second semester three students were put in the basic first-grade arithmetic text and others almost completed the second-grade basic text.

The students also had many opportunities to widen their experiences and vocabulary through field trips to the Milwaukee zoo, Racine Public Library, Mitchell Conservatory in Milwaukee, and Racine Historical Museum. These trips served as motivation for lessons in the language arts, where developmental activities of the Unit were focused. A description of the language program follows.

Members of the Unit were concerned about the oral language development of their students, most of whom came from a culturally disadvantaged home environment. The intent was to provide language readiness experiences. The specific problem to be studied was the possibility of bringing about a significant improvement in language achievement by providing supplementary language experiences.

Subjects

Subjects in the experiment were the 71 second-grade students at Franklin School. About 70 percent were Negro and 30 percent Spanish-American. Ages ranged from 7.0 to 9.0 years at the start of the experiment, with the mean age being 7.8.

Method and Procedure

Subjects were stratified on the basis of sex and assigned to one of four groups. One group of 14 subjects was the control group which received a health unit in place of the language experiences received by the other three groups.

The three experimental groups all received the same treatments. Subjects in these groups received a variety of language experiences. The primary emphasis was on oral experiences such as choral speaking, creative dramatics, play reading, etc. These subjects also received instruction in written language beginning with very simple sentence writing and progressing to longer, more creative compositions.

Throughout all the language experiences indirect attempts were made to improve the self-image of experimental subjects. This was done by selecting particular topics and materials to be used in the language program.

Each group met for three one-half hour periods each week, for a period of twenty weeks. The three experimental groups worked as separate groups part of the time and in a single large group part of the time, with special individual or small-group instruction for some students. The teachers rotated among the four groups on a weekly basis. One teacher was free each week, to assume primary responsibility for planning the experimental activities for the following week.

Data Collected

Pre- and postexperimental writing samples were obtained from all students. All samples were then rated by the teachers on a 5-point scale with respect to capitalization, punctuation, presence of a complete thought, and pertinence to the assigned topic; analysis was based on the total of these four ratings, making the maximum possible score twenty. Additionally, the number of words written was recorded and raters judged the amount of improvement shown, also on a 5-point scale.

Sociometric data were also collected before and after the experiment. Analysis of this variable was based on the number of times each child was chosen within his homeroom on the two occasions. Therefore, the maximum score possible for a child was twice the number of children in the classroom.

Additionally, an attitude questionnaire was given on the two occasions. Possible scores ranged from 0 to 40. Group means for all these measures are presented in Table 26.

Table 26

Group Means for Measures
Used in this Project

Variable	Groups	
	Experimental	Control
Writing Sample: Pre	6.98	6.93
Writing Sample: Post	12.49	12.93
No. of Words: Pre	16.53	16.93
No. of Words: Post	26.66	23.93
Sociogram: Pre	2.13	1.66
Sociogram: Post	2.15	2.72
Attitude Scale: Pre	33.51	37.51
Attitude Scale: Post	36.43	38.22
Rated Improvement	3.43	3.43

Another measure which was obtained at the beginning and again at the end of the experiment was an evaluation of a tape-recorded oral presentation by each student. Three teachers, including the Learning Specialist, rated each presentation on a five-point scale. A child's rating then was the mean of these three ratings. No formal analysis was made of this measure, but mean ratings for both groups on both occasions were calculated. These are presented in Table 27.

Table 27

Mean Ratings of Oral Presentations

Groups	Pre	Post	Gain
Experimental	3.2	3.4	.2
Control	3.1	3.6	.5

Results and Discussion

All data except the oral evaluation were analyzed using a 2 x 2 analysis of variance or covariance. None of these analyses revealed significant differences between treatment and control groups. Significant sex differences, favoring females, were found for the writing sample ratings, and for the number of words written for the sample.

It is possible, of course, that many activities of the type intended for the language enrichment treatment, were used in the group being taught the health unit, thus obscuring any real effect due to additional language experiences. Then, too, experiences in a language enrichment class constitute such a small portion of a child's daily written and oral expression, that strong effects over a relatively short period of time cannot reasonably be expected. The possibility that the measures used were insensitive to real differences between the groups also exists. While no significant treatment effects can be reported, then, the language enrichment program conducted over a long period of time is still considered to have possibility for contributing to improved performance.

LANGUAGE ARTS FOR DISADVANTAGED CHILDREN, FRANKLIN SCHOOL, KINDERGARTEN

Experiences with children from poor home backgrounds led the members of the S & I Unit—the Learning Specialist, Dawn Hafton, and teachers Diana Fish, Jeanne Keenan and Judy Larson—to decide that such children might profit from intensive work in developing language and communication skills. Dr. Glenn Dagnon, Dean Emeritus of the S & I Center, played an active role in much of the planning and implementing of the research design. In cooperation with Elizabeth Williams and Marcella O'Leary, Language Arts and Music Consultants for the Racine Unified School District, the resulting personal development program to provide an "overdose" of experiences in each of three areas: vocabulary development (8 words), music (8 words), and language development (18 words).

In order to determine the effectiveness of this program an experiment was designed to investigate the following hypotheses: first, that an overdose of language experiences, in place of the usual play-time activities, would facilitate the development of language and communication skills in culturally disadvantaged kindergarten children; and secondly, that such a program would be more effective when used in small-group than in large-group instruction.

Subjects

Subjects for the experiment were the 116 children in the kindergarten S & I Unit at Franklin School, Racine, Wis. They ranged in age from 5.25 to 7.25 years with the average being 5.67 years. Approximately 75 percent came from families which would be classified as "lower class." About one-half were Negro and one-fourth Spanish-American. For the experiment no distinction was made between children in the morning kindergarten and those in the afternoon.

Design

A randomized block design was employed to assign students to treatments, with blocking on aptitude, as measured by the Quick Test, and sex. Three levels of aptitude were employed, a total of eight groups were used in the experiment: three small ($n=8$) groups, three large ($N=50$) groups, and two control groups (1 and 15 yr). For purposes of analysis, 1y to small groups were combined, as were those in large groups and control groups. Thus, for the final analysis, there were 16 1y who had received small-group instruction (experimental), 60 who had received large-group instruction (experimental), and 17 in the control group.

Procedure

Well before lessons were presented, each of children in the six experimental groups received the same treatment, although the groups were different.

The main objective of these "overtures" sessions was to acquaint these children with many first-hand experiences about which they were encouraged to verbalize, dramatize, and sing. For example, the Peabody Vocabulary Kit was often used to stimulate interest and enthusiasm. The picture in this helped the children identify community helpers such as playground attendant, milkman, nurse, and doctor. The children were first acquainted with the names by holding up the cards and labeling. Then a riddle game was played where the teacher said, "I have some riddles about what these people do." The children would respond to the various riddles by holding up the card and naming at the same time.

Demonstrations then alternated the children to act out various ways these people help the community. Sometimes children were given an opportunity to don a policeman's hat or a pair of firemen's boots to make their dramatizations seem "real."

Live models were brought into the classroom to help these children identify with community adults and also to visit and play with them. A well known Negro chef dressed in his chef's attire made a lasting impression on these children. In addition to talking to them about his job, he talked with them about such things as the importance of cleanliness and neatness. After he demonstrated making pumpkin pie, the children learned that's how and made their own pumpkin pies. The chef circulated among them verbalizing and singing with them. "This is the way we make a pie, roll the dough," etc.

These children often had the opportunity of hearing themselves. Tantine songs, finger plays, stories, and telephone talk were often taped and played back for the children to hear. During the last eight weeks of the experiment much emphasis was placed on "creative activities." In one lesson on radio-discrimination, a tape recording was played of sounds from home and then sounds from school. The children did very well in identifying sounds as sounds were similar for instance, for the egg beater sometimes said have mixer. Other sounds used were the telephone ringing, dishes the telephone, a broom sweeping and closing, an alarm clock ticking, a lawn sprinkler, water running, an school the sounds included children coming down stairs, piano, musical instruments, cutting paper, and door closing. The second lesson included identifying the teachers' voices. A tape was played with the four teachers each saying a line of a nursery rhyme. Then two of four children were chosen to go behind a divider and say a nursery rhyme while the teacher stood them. The other children then tried to identify which child was speaking.

Another lesson was on the visual discrimination of the color words. Large circles with a color painted around the edge and the name of that color in the center were placed on the floor in the center of the room. Each child was given a card with a color name on it. He would walk around and match the color word to the correct circle. This was a good game, as the children could move around and also help one another. In the small groups the children could have more than one turn, and it became a contest with much interest shown.

The two teachers presented the lessons cooperatively, and teachers rotated among the different groups, in a pre-established sequence (table square), in order to avoid contamination of experimenter and teacher variance.

A teacher-made test of conceptual ability was administered twice during the experiment, once at midyear and again at the end of the experiment in May. The test was used to measure the children's vocabulary and comprehension. The children were tested individually. Each child was shown a card containing four simple pictures. The test administrator would then say a word and ask the child to point to the appropriate picture for the word. Thirty-six cards were used. An item analysis of the first administration of the test revealed that it had an internal consistency reliability of .81, with scores ranging from 2 to 32 and a mean of 20. Only two items in the test (number 1 and 18) failed to discriminate adequately, both because they were too easy.

The analysis was based, then, on the design in Table 1.

Table 10

Meanings of Language Enrichment Experiences, Franklin School

Groups	Ability	Sex	Measures 1	2
Small (S)	High	M	A = B	
		F	A = B	
	Average	M	A = B	
		F	A = B	
	Low	M	A = B	
		F	A = B	
Large (L)	High	M	A = 1.0	
		F	A = 1.0	
	Average	M	A = 1.0	
		F	A = 1.0	
	Low	M	A =	B
		F	A =	B
Control	High	M	A = B	
		F	A = B	
	Average	M	A = B	
		F	A = B	
	Low	M	A = B	
		F	A = B	

Because of the wide variation in cell sizes, an unweighted-means analysis of variance was used.

Results

Table 10 presents a summary of the analysis of variance of scores on the Stanford-Binet. The analysis indicates several things quite clearly: first, the differences are virtually nonexistent. (Note that the F ratios for the two main effect and all interactions with this variable are all less than 1.00.) Secondly, there is no interaction between Group and Ability, or between these two variables and either of the others.

There is a significant difference between groups (Table 9), and the means for the three groups are presented in Table 11. Two orthogonal comparisons had been planned for this factor, one to compare the combined treatment groups with the control, the other to compare the small with the large group. For the first comparison, the resulting F ratio was 7.11, significant beyond the .01 level. The second comparison yielded a nonsignificant F of 1.00. Thus, the two experimental groups did not differ significantly from one another, but were significantly superior to the control group.

Table 9

Summary of Analysis of Variance

Source	df	MS	F
Between Subjects 119			
Groups (A)	2	677.04	13.34**
Ability level (B)	2	648.33	12.80***
Sex (C)	1	31.50	<1
A × B	4	33.10	1.00
A × C	2	6.37	<1
B × C	2	0.08	<1
A × B × C	4	0.50	<1
Ss within ABC	108	50.11	
Within Subjects 120			
Measures (D)	1	1815.15	36.67***
A × D	2	52.08	1.05
B × D	2	38.31	1.00*
C × D	1	0.70	<1
A × B × D	4	2.31	1.00
A × C × D	2	1.74	<1
B × C × D	2	1.38	<1
A × B × C × D	4	0.83	<1
D × Ss within ABC	108	2.03	
Total	239		

* p < .05
** p < .01
*** p < .001

Table 11

Means for "Groups" Factor

	Groups		
	Small (S)	Large (L)	Control
Means	111.10	108.50	98.37

The ability factor represented a highly significant (p < .001) source of variation. The means for the three ability levels are presented in Table 12. In this case, performance of students in the low ability group was significantly lower than that of pupils in the average and high groups, but these latter two groups did not differ significantly from one another (p > .05).

Table 13 presents the means for the significant Groups × Measures interaction. The interpretation of this interaction is relatively straight-

Table 8A

Means for "Ability" Factor

	Ability Level		
	High	Average	Low
Means	30.62	27.38	24.12

Table 8B

Means for Groups × Measures Interaction

	Measures	
Groups	One	Two
Small (E)	27.81	30.00
Large (E)	26.71	28.00
Control	25.33	27.10

forward, reflecting the fact that the two experimental groups showed more improvement from the first to the second measure than did the control group. Students receiving small-group instruction also showed more improvement than those receiving large-group instruction.

A somewhat different situation is reflected in Table 8C, which contains the means for the significant Ability Level × Measures interaction. In this case, there is an inverse relationship between ability and improvement.

Table 8C

Means for Ability × Measures Interaction

	Measures	
Ability Level	One	Two
High	28.14	31.71
Average	26.57	28.21
Low	23.57	24.68

Discussion

The significant Group effect, resulting from the superiority of students in the treatment groups over those in the control group, provides convincing evidence that a carefully designed program to provide language and communication experiences can effectively improve the conceptual abilities of children at the kindergarten level. Although students receiving the instruction in small groups performed better than those in large groups, this difference was not significant. The interaction is, then, that while small-group instruction may be beneficial, its effects are not nearly as important as is participation in a well designed enrichment program.

The significant differences resulting from ability are just what would be expected and do not reflect a unique or particularly important finding of this experiment. This conclusion also applies to the significant effect on the "Measures" dimension: that is, it is reasonable to expect improvements in conceptual ability over time, as a result of such factors as maturation and history, as these terms are used by Campbell and Stanley.[8]

What is important, however, is the interaction of Group and Measures. Figure 1 indicates that students in the experimental groups showed more improvement over the three months between the two measurements than did those in the control group. If we assume that change is a linear function, and extrapolate back to three months before the first testing (just after the beginning of the experiment), the means for the three groups would be 27.50 (small group, experimental), 27.14 (large group, experimental), and 25.23 (control group). Interestingly, no manners was taken at that time, so that these values are hypothetical. It seems reasonable, particularly when one considers the close agreement of these values, that the hypothetical approximates the real to this situation.

The Ability × Measures interaction reflects the greater improvement of the low ability group. There are two possible sources for this effect. First, the experimental program might have been more effective with students with low scores on the Quick Test, because of the fact that these do tend the greatest need for the kinds of enrichment activities employed. On the other hand, the difference may reflect a "ceiling effect" of the test employed. Support for this second explanation comes from the fact that approximately 12 percent of pupils in the high and average ability groups attained scores of 32 or 33 out of a possible 33, as contrasted to fewer than seven percent of pupils in the low ability group. The most probable explana-

[8] Campbell, Donald T., and Stanley, Julian C., "Experimental and Quasi-Experimental Designs for Research on Teaching" in Gage, N. L., editor, Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

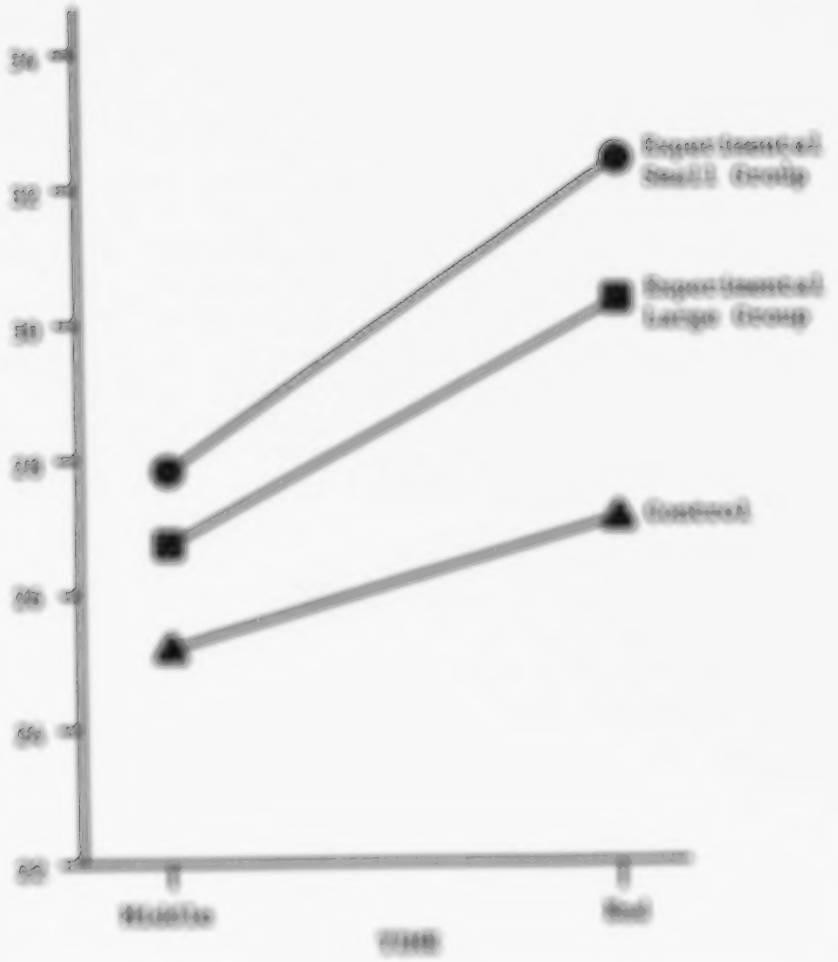

Figure 1. Performance of Treatment Groups on Teacher-Constructed Test at Midterm and Conclusion of Experiment, Franklin School, Kindergarten

[illegible] in the control groups were losing background of some good experiences. One teacher felt that kindergarten children need the security of the same teacher. Another teacher felt that the children's opportunity to know more than one teacher was a motivational inspiration in itself and that the children looked forward to having different teachers.

Implications

On the basis of these findings, the obvious conclusion is that a program such as was carried out in this experiment has the highly desirable effect of improving language and communication skills. Although a systematic follow-up to determine the extent to which these skills influence subsequent scholastic performance is needed, the evidence presented here provides a strong argument for the extension and implementation of such a program in kindergartens of inner city schools.

Since language is a prerequisite to almost all learning, helping children to communicate easily and understandably is one of the big responsibilities of preschool, kindergarten, and first grade.

Research carried on throughout the country indicates that language development programs are often unsatisfactory and incidental. Teachers consider do not capitalize enough on the incidental opportunities to stimulate children's [illegible]. Often they say they do do not have a listening ear. Children should be encouraged to respond by creative [illegible], systematically stimulating to increase verbal output, by providing opportunities for [illegible], and consideration.

Teachers in the preschool and kindergarten have felt the need for the child to have many experiences to expand his world. Observation games and devices, much from our collection also, and the creative arts have been the setting for most programs. Perhaps what is needed is for teachers to define more specific goals for the classroom "experiences" which would include the most effective way to present the motivational techniques needed to elicit the child's response, and the sincere effort on the part of the teachers to provide feedback and reinforcement.

A language development program should probably include activities to develop both receptive and receptive language. Expressive language is the child's way of verbally expressing himself and is important as a means of communication. Receptive language develops understanding and comprehension. Both are vitally important for achievement in reading.

6

[redacted]

[redacted]

The fifth-grade Unit at Hewett School, Glen Humphrey, principal, consisted of Jerome Sullivan, learning specialist, and teachers Joe Bee and Al Heywood. Each staff member had a room of station and these rooms were adjacent. Formally, pupils and teachers met in more than one room, and teachers met with more than one group. A half-time teacher aide assisted the certified staff in working with 60 students.

The Unit staff desired to reduce the social conflict among the student population by helping parents to cope with problems which their children brought to them from school, and by creating a climate for the pupil which would reduce the frustration of learning and improve academic achievement. To this end, a variety of activities were planned to involve all parents of children in the Unit. These included a get-out-the-vote, boys' recreation night with fathers, girls' recreation night with mothers, girls' recreation night with mothers, and fund raising projects involving parents and children, such as dinners eats, bake sale and a skate drive. Parents assisted the instructional staff by giving talks about their jobs and places of employment, arranging field trips to factories where they worked, and serving as guides or chaperones on the trends and field trips taken by the Unit.

An experiment was designed to determine whether greater contact between the home and school would improve students' academic performance and parents' perception of the school. The experiment was carried on from October, 1972, through May, 1973.

Subjects

Subjects for this experiment were the 57 fifth graders at Hewett School and their parents. The student population comes from a blue-collar community.

Treatment

Subjects were randomly assigned to the two treatments. A brief description of the treatments follows.

Experimental. The first group of parents were visited in their homes, asked to be part of a parent study group, given assistance with their child's homework, given an opportunity to participate in school activities, and had a teacher-Unit with each report card.

The contact between the home and the school was primarily undertaken by the Unit leader. Those assisting with the Unit were a home-school parent from the community, a social worker, teachers of the Unit, and the principal. The teachers and the Unit leader held a conference with each of the parents of the experimental group. These parents also received home visits during the and immediate notification in case of an outstanding learning or behavior change. This was accomplished by phone call, home visit, or personal note.

Control. The second group of parents were contacted by the school in accordance with normal procedure.

Data Collected

A 20 item questionnaire was constructed by the school personnel. Some of the questions were related to the desirability of school-home contact for particular situations and to home and school factors influencing student performance. The instrument was administered in October and in May to the parents of all pupils enrolled in either treatment.

The questionnaires were delivered to the kit to each home by the Unit leader, while in the spring it was sent home with the pupils. Table 10 indicates the return rate of the instrument.

Additional data used to evaluate the success of this program included field test data from fall and spring administrations of the Stanford Achievement Test.

Table 10

Return Rate of Parent Questionnaire, Stowell School, Grade 5

		Caterpillar	Wasp
Experimental	Returned	17	20
	Not returned	11	2
	Total	28	22
Control	Returned	17	21
	Not returned	10	8
	Total	27	29

Analysis of Data

Responses to the questionnaires administered in the fall were subjected to a scaling analysis by means of the reciprocal averages technique. This procedure assigns item responses to such a manner that internal consistency is maximized. The reliability of this instrument agreed with the results as presented ($r = .83$).

Weights obtained from analysis of the fall questionnaires were used to obtain total scores for both fall and spring administrations. An analysis of covariance was then performed on total spring scores of parents who responded to the questionnaire on both occasions. This restriction reduced the number of subjects for the analysis to 17 and 18, respectively, for the experimental and control groups.

The results of the analysis of covariance are found in Table 11. The significant treatment effect represents the fact that parents in the experimental group had significantly higher scores than those in the control group despite the fact that the control parents had initially higher scores, as indicated in Figure 5.

A multivariate analysis of covariance was performed on the ten subtests of the Stanford Achievement Test using (1) as a covariate. Results of the analysis indicated that the experimental students performed significantly better than did the control group subjects ($p < .05$) when the data were adjusted for the higher IQ of the control group. There were substantial differences between the adjusted scores of the two groups in language, arithmetic

Table 11

Analysis of Covariance Performed on Parent Questionnaires

Source	SS	df	MS	F
Treatment	2858.2111	1	2858.2111	18.88***
Error	8455.3418	28	151.3454	
Total	8771.8.1.69	29		

*** $p < .001$

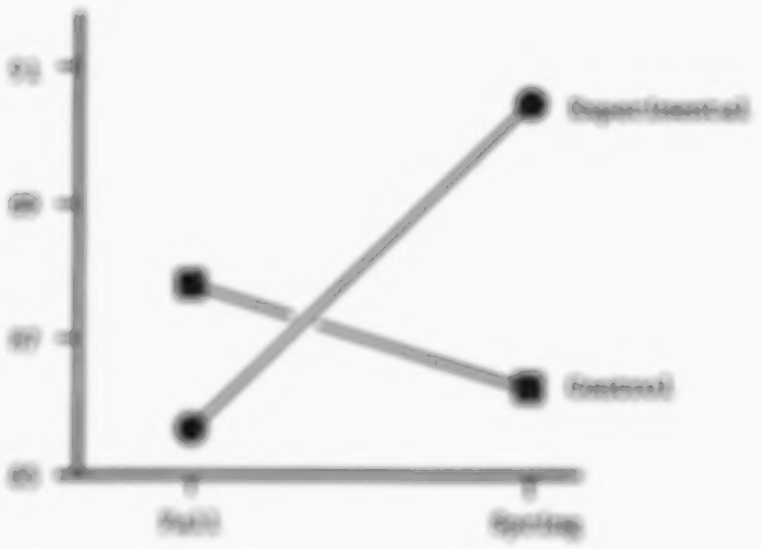

Figure 5. Means of Treatment Groups on Fall and Spring Administration of Attitude Questionnaire

computation, arithmetic applications and science. All groups of the experimental group. In Figure 5 grades of the same grade equivalents (unadjusted) for each group are presented, although the experimental group's performance was slightly lower than that of the control group on the five remaining subtests, the scores adjusted for IQ, as were those scores, generally favored the experimental group.

Discussion

The experimental treatment seems to have been effective both in motivating student performance and in making parents' attitudes toward the school more positive. It is possible, as well, that such a program had other unanticipated benefits, such as positively changing student attitudes. The effects of any public relations program, of course, may also extend to performance of other pupils in the family. That such a time-consuming program was well worth the effort of the many community personnel who contributed is evident.

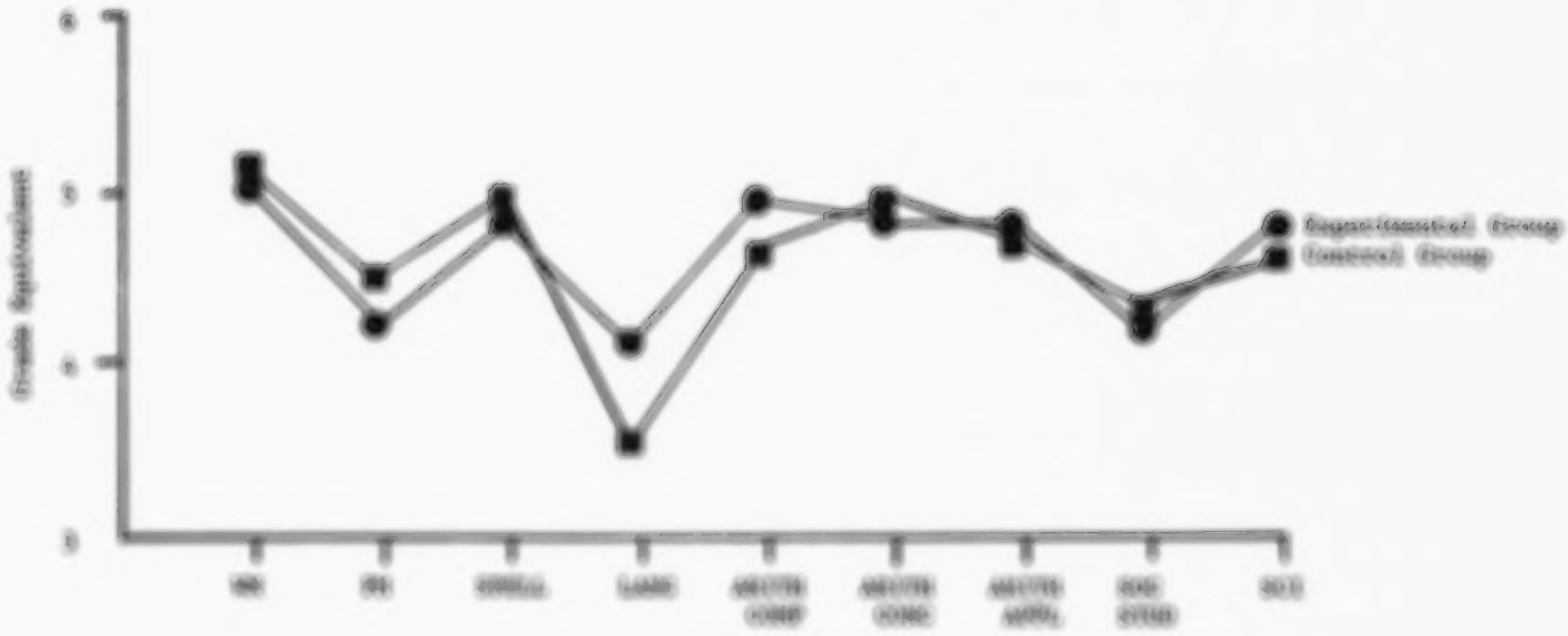

Figure 1
Profile of Means on Stanford Achievement Test for Experimental and Control Groups, Grade 2, Hewett School, Spring 1997

4

FIELD TESTING A K-1 UNIT, RACINE UNIFIED SCHOOL DISTRICT (1966-67)

A preliminary plan for field testing in Racine was outlined in Working Paper No. 1 (Wirtanen et al., 1966).[5] A central feature of this plan was to use regression analysis for Grades 1 through 3. Three other essential features of the plan were (1) that children in control and K-1 Units should be as alike as the same distributions of IQ level, educational achievement level, socioeconomic level, and interest in schooling as shown in attendance records; (2) that the instructional program of the two schools should be alike in terms of amount of time spent on the various curriculum areas and should be described carefully on the basis of observations could be clearly established; and (3) that the standardized tests should be administered by the same personnel but not by the Unit teachers and teachers directly involved with the instructional program in either the K-1 Units or the control schools.

A regression analysis was not possible because adequate pretest data were not available from all groups. Also, as is often the case between business conditions and school activities, several essential conditions related to comparability of control school classes, their instructional programs, and standardized test administration were not met. For the four K-1 Units and their controls from which complete baseline data (IQ and a standardized achievement test) were available, analysis showed, as all but one group, numerous significant differences. In the absence of both randomization and initial comparability, analysis of covariance, another strategy presented in Working Paper No. 2, was also deemed inappropriate. Therefore, the most meaningful information about the effectiveness of the Units must be in terms of gains made on standardized tests by the Unit children, relationships in terms of consistency with children and programs in control schools. In other words, comparison of the performance of K-1 children with the large national test standardization population of children. Gains can be identified as the relevant test manual as more useful than comparison with a Racine group of children that is very different from the K-1 group or that may have received a quite unique instructional program.

Such scores, of course, may be calculated only when tests have been administered to pupils on two occasions. Such adequate baseline data were not gathered from pupils at the kindergarten or first grade, and because the responsibility of K-1 and control groups is to be questioned, no analysis of the criterion data for these levels were performed.

The remainder of this field testing report, then, deals with K-1 Units and their controls in Grades 2, 3, and 4. The control schools were selected early in the fall of 1966 by personnel of the Racine Unified School District with the intention of securing control groups reasonably similar to K-1 Unit schools. The schools of the K-1 Units and their control schools and the data gathered in them are presented in Table 10. Baseline data were gathered in October, 1966, and criterion data in April, 1967.

THE SECOND-GRADE UNIT AT HORLICK SCHOOL

Baseline and criterion data were gathered for 27 pupils in the Horlick K-1 Unit and 28 pupils in the control school. The mean IQ of the K-1 pupils was 101.2, that of the control

[5] Wirtanen, James L., Fagnan, Glenn E., Glennon, Vincent J., Gunnell, Barbara L., and Cook, Doris M. A plan for field testing K-1 Units. Working Paper from the Research and Development Center for Cognitive Learning, University of Wisconsin, 1967, No. 1.

Table 10

Locations, Levels, and Instruments for Field Testing in Reading

Location of Q & I Unit	Location of Control Group	Grade Level	Baseline Tests	Criterion Tests
Stovall	Winslow	5	Gilliland-Anderson (G) Standard Achievement Test, Primary II	Standard Achievement Test, Primary II
Glenn	Roosevelt	8	Gilliland-Anderson (G) Standard Achievement Test, Intermediate II	Standard Achievement Test, Intermediate II
Stovall	G. Hall	6	Gilliland-Anderson (G) Standard Achievement Test, Intermediate II	Standard Achievement Test, Intermediate II
Winslow	Franklin	5	Gilliland-Anderson (G) Standard Achievement Test, Intermediate II	Standard Achievement Test, Intermediate II

grade (10.8). Despite the similarity of mean ability, mean achievement scores obtained in the fall upon administration of the Standard Achievement Test, Primary II battery vertically favor the control school as figure 5 indicates.

The median achievement across subtests of the Q & I Unit was 1.6. That of the control school pupils, 1.8. The control school pupils exceeded that of the Q & I school on every subtest with the average difference being .2 years. All of these differences were statistically significant. That such pronounced differences in achievement were also found to be consistent further justifies the decision to eliminate consideration between Q & I and control school pupils' previous achievement test data as unavailable.

Criterion data were analyzed using 2-way analyses of covariance, with the factors being Group (Q & I Unit vs. Control) and Sex (Male vs. Female). One covariate was employed for each analysis. One of these was IQ, the other the subtest of the Standard Test battery which corresponded to the subtest from which testing. Table 11 summarizes these analyses. As this table indicates, the only significant effect was the Group effect on the Spelling analysis. In this case, the students in the Q & I Unit averaged slightly more than one grade better than the control group, after initial differences were removed. This represents an improvement of 1.3 years (grade equivalent) in the Q & I Unit, as contrasted with 1.0 years in the control group.

The pre- and posttest grade equivalent scores and gains for both groups are presented in Table 12.

We find that the gain of students in the Q & I Unit is at minimum one and one-half times that of pupils in the control school. The median gain in the six-month interval between and administrations across subtests was .8 years for the Q & I pupils, .4 years for the control pupils.

The Q & I Unit, then, compared to control in every subject area. Furthermore, whereas the typical Q & I pupil's performance was 0 to 8 months below grade level at the time of fall testing (1.8), the performance was only 1 to 8 months below grade level on the spring test. For this group, then, the one between some performance and national norms narrowed substantially. We may conclude that the Q & I Unit performed its instructional function well.

The Fourth-Grade Unit at Cove School

Analyses of baseline data revealed significant differences favoring the control group at Roosevelt School on (1) and the following subtests of the Standard Achievement Test: Spelling, Language, Arithmetic Computation, Arithmetic Concepts, and Arithmetic Problem Solving. The gains (1) for the Q & I Unit was (10.8 and for the control was (10.5.). Figure 7 portrays differential performance of these and control school pupils on the Standard Achievement subtests. Control school pupils outperformed Q & I pupils on all subtests, with the median difference being .2 years.

Because baseline data given no support to the notion that these schools were comparable, further comparative analyses were not performed.

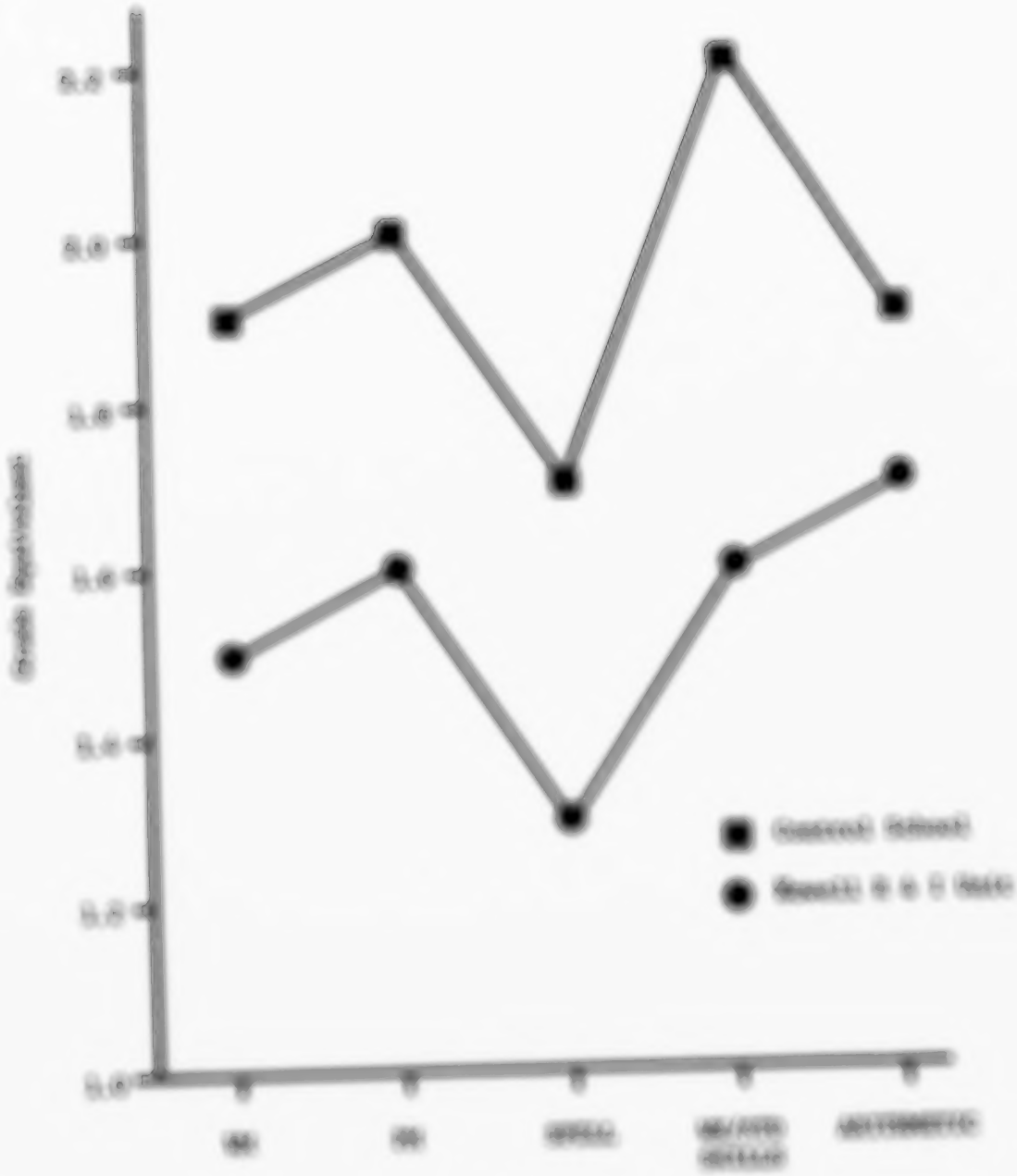

Figure 5.
Mean Performance of Normal and Control School Pupils,
Creative Environment Test, Oral, 1966

Table 37

Summary of Analysis of Covariance: Stowell School, Grade 6

Source of Variation	df	Dependent Variable: Word Meaning MS	F	Paragraph Meaning MS	F	Spelling MS	F
Group	1	6.65	<1	60.28	<1	182.68	5.70*
Sex	1	38.88	5.55	55.87	<1	5.17	<1
G x S	1	59.12	1.65	51.87	<1	8.11	<1
Covariation	3	1818.78	188.21***	2868.18	108.81***	2628.30	135.18***
Error	137	17.18	----	26.61	----	17.96	----

Source of Variation	df	Dependent Variable: Word Study Skills MS	F	Arithmetic Computation MS	F
Group	1	73.18	<1	8.18	<1
Sex	1	135.58	1.79	1.71	<1
G x S	1	1.88	<1	2.88	<1
Covariation	3	6808.18	86.88***	1215.26	65.88***
Error	137	78.78	----	18.16	----

* $p < .05$
*** $p < .001$

Table 38

Stowell and Control Schools, Grade 5, Fall and Spring Group Means on Stanford Achievement Test

	Q & I			Control		
	Fall	Spring	Gain	Fall	Spring	Gain
$\bar{X}_1$ Word Meaning	1.58	2.18	.8	1.91	2.45	.5
$\bar{X}_2$ Paragraph Meaning	1.88	2.18	.8	1.96	2.42	.4
$\bar{X}_3$ Spelling 5	1.18	2.18	1.1	1.78	2.62	.9
$\bar{X}_4$ Word Study Skills 5	1.82	2.48	.8	2.18	2.65	.5
$\bar{X}_5$ Average Arithmetic	1.87	2.17	.5	1.96	2.27	.3

schools were one to two years below grade level in performance across subtests. Fall subtest means favored Stowell pupils two times and control pupils four times.

Analyses of covariance of criterion data were performed, using IQ and the concomitant baseline subtest as covariates for each criterion subtest. Table 37 summarizes these analyses. Of the four significant Group effects shown in Table 37, one—for the Social Studies subtest—favors students in the control group. The other three—on the Word Meaning, Arithmetic Concepts, and Arithmetic Applications—resulted from the superiority of students in the Q & I Units.

Inspection of Table 41 will reveal that something very encouraging has happened in this Unit. Baseline scores of this group

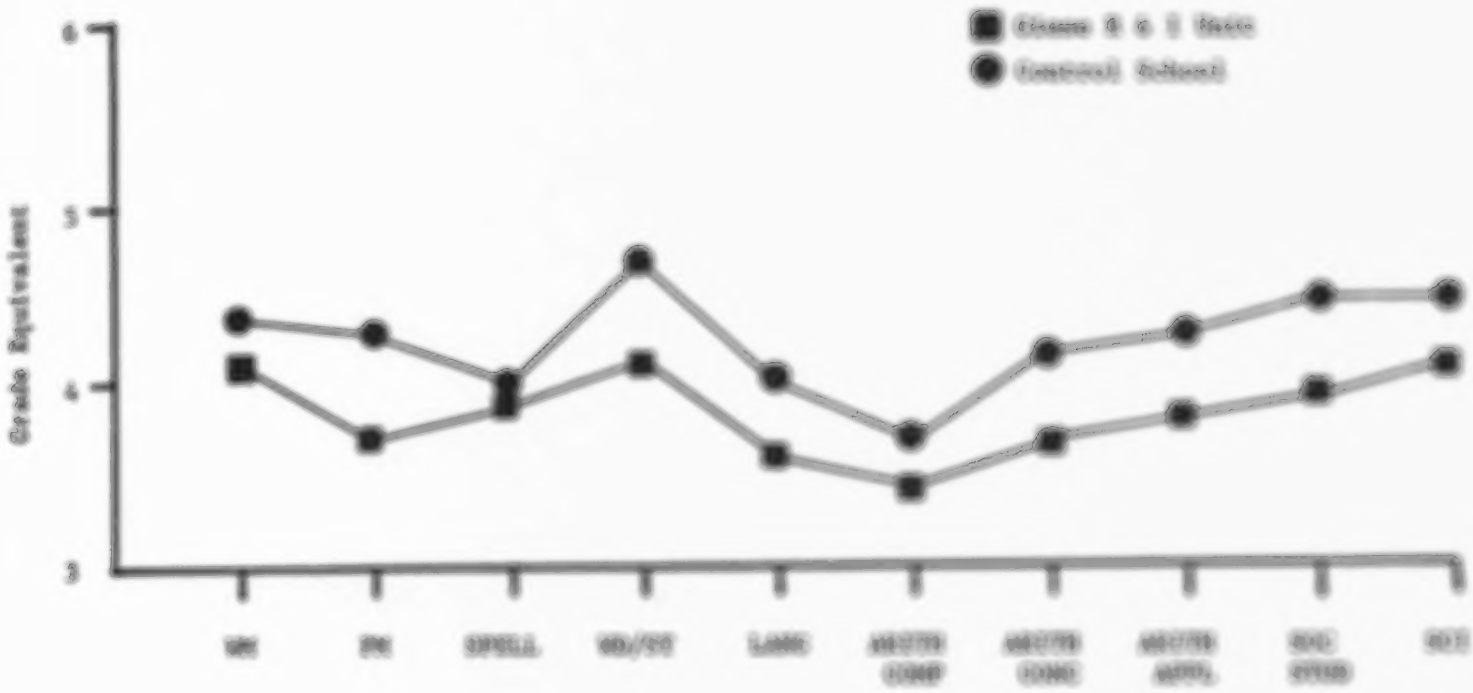

Figure 5.

Mean Performance of Grade and Control School Pupils, Stanford Achievement Test, Fall, 1968

Table 28

Grade School, Grade 4, Fall and Spring, Group Means on Stanford Achievement Test

Variable	Fall	Spring	Gain*
Word Meaning	4.2	4.8	1.6
Paragraph Meaning	3.7	4.5	1.8
Spelling	3.9	4.6	1.7
Word Study Skills	4.2	4.6	1.4
Language	3.6	4.3	1.7
Arithmetic Computation	3.5	4.1	1.6
Arithmetic Concepts	3.7	4.5	1.8
Arithmetic Application	3.8	4.4	1.6
Science	4.2	4.8	1.6
Social Studies	4.0	4.7	1.7

*Expected gain, based on published norms and interval between testings, was 0.5 years.

averaged about 1.35 years below grade placement, indicating (if we assume linear changes in performance) that these students have gained about 1.35 years less over time than the test norms indicate is to be expected. During the 5 month interval between the two administrations of the test in Grade 5, however, the median gain for R & I pupils was 7 months, while for the control school it was 5 months.

Of course, data are not available to support the hypothesis of a cumulative linear decrement in performance through the first four grades, but even if some other pattern were found, these results would still be highly encouraging. It should also be noted that performance was quite uneven, with the greatest improvement (0.8 years, grade equivalent) being shown on the Word Meaning subtest, and the least improvement (0.2 years) being shown on the Social Studies subtest.

THE FIFTH-GRADE UNIT AT WINSLOW SCHOOL

Analyses of baseline data revealed that there were significant ($p < .001$) differences favoring the R & I Unit on all measures. Table 42 presents the means for the two groups on these measures.

Table 42 makes it obvious that the "control" group was not at all appropriate. Students in

Table XI

Summary of Analysis of Covariance: Howell School, Grade 5

Source of Variation	df	Dependent Variable: Word Meaning MS	F	Paragraph Meaning MS	F	Spelling MS	F
Group	1	1,030.08	7.22**	58.38	<1	121.89	<1
Sex	1	6.11	<1	1,088.58	8.12**	830.37	5.88
G × S	1	0.89	<1	68.01	1.75	7.18	<1
Covariation	1	1,87[illegible]5.97	715.80***	1,098.79	28.78***	1,098.59	778.11***
Error	130	18.30	----	515.62	----	28.30	----

Source of Variation	df	Dependent Variable: Language MS	F	Arithmetic Computation MS	F	Arithmetic Concepts MS	F
Group	1	1.48	<1	17.04	1.04	38.18	4.11*
Sex	1	880.08	6.11**	5.08	<1	5.88	<1
G × S	1	188.18	1.30	4.20	<1	4.75	<1
Covariation	1	8[illegible]78.60	1,60[illegible]***	918.04	56.81***	888.75	1,[illegible]***
Error	130	87.87	----	16.18	----	9.39	----

Source of Variation	df	Dependent Variable: Arithmetic Applications MS	F	Social Studies MS	F	Science MS	F
Group	1	81.87	5.28**	1,[illegible]78	5.28**	88.18	5.18
Sex	1	1.45	<1	1.82	<1	2.19	<1
G × S	1	1.82	<1	85.18	5.18	18.15	1.19
Covariation	1	78[illegible]7.08	518.79***	1,[illegible]11	88.77***	1,[illegible]1	81.18***
Error	130	15.24	----	16.08	----	15.88	----

* p < .10
** p < .01
*** p < .001

the control group averaged 1.7 years below those in the R & I Unit in the fall of 1968. The 1.5 point difference in mean IQ for the two groups is another indication that they were not comparable. Therefore, further comparisons between the two groups were not made.

Table 12 presents mean grade equivalent scores for R & I Unit students from the fall and spring administrations of the Stanford, and the gains during the interval.

Table 13 indicates that the median subtest gain was seven months in the five month interval between test administrations. By the end of the year students were performing, on the average, below grade level in only one area, arithmetic computation, and had made growth even in that area consistent with their past performance. The median grade equivalent across subtests was 6.1 in April of the fifth grade compared with 5.3 in October. As was noted before, unusually large gains were made in some of the language arts, in which the developmental activities of the Unit staff were focused.

Since scores R & I Unit-control group differences at the beginning of the year made comparisons of the two virtually meaningless, evaluation of the Unit in terms of gain scores seems warranted. These gains presented demonstrate that for most subject areas student growth was at least as great as it had been for this group in years past, and that for some subject areas the growth was greater than would normally be expected.

Table 51

Standard Achievement Test Scores for Hawaii Grade 1 and Control School, Fall, 1986, and Spring, 1987

	Word Meaning	Date	Paragraph Meaning	Date	Spelling	Date		Date	Lang.	Date
Hawaii										
Fall	1.8[E]	1.2[E]	1.5[E]	.8	1.5[E]	.8			1.5[E]	.8
Spring	2.0[E]		2.3[E]		2.3[E]				2.3[E]	
Control										
Fall	1.9[E]	.8	1.8[E]	.6	1.5[E]	.8			1.5[E]	.6
Spring	2.7[E]		2.4[E]		2.3[E]				2.1[E]	

	Arithmetic Computation	Date	Arithmetic Concepts	Date	Arithmetic Application	Date	Social Studies	Date	Science	Date
Hawaii										
Fall	1.1[E]	.8	1.3[E]	1.0[E]	1.3[E]	.8	1.1[E]	.8	1.3[E]	.6
Spring	1.9[E]		2.3[E]		2.1[E]		1.9[E]		2.1[E]	
Control										
Fall	1.3[E]	.8	1.7[E]	.8	1.3[E]	.8	1.3[E]	.6[E]	1.3[E]	.5
Spring	2.1[E]		2.5[E]		2.3[E]		2.3[E]		2.3[E]	

[E] = Fall, Hawaii significantly superior.
[E] = Control group significantly superior.

Table 52

Means[a] for Groups on Stanford Measures

Variable	Fall, Hawaii	Control
IQ	108.9	99.9
Stanford Word Meaning	1.1	1.2
Stanford Paragraph Meaning	1.3	1.3
Stanford Spelling	1.4	1.4
Stanford Language	1.3	1.4
Stanford Arithmetic Computation	1.4	1.1
Stanford Arithmetic Concepts	1.4	1.4
Stanford Arithmetic Applications	1.3	1.3
Stanford Social Studies	1.4	1.1
Stanford Science	1.4	1.4

[a]Except for IQ, the means are expressed in grade equivalents.

Table 53

Contrast Means and Ratios on Standard Achievement Subtests, Winston School, Grade 1[a]

Variable	Spring '87	Date
Word Meaning	2.0	1.1
Paragraph Meaning	1.8	1.4
Spelling	1.1	1.2
Language	1.8	1.4
Arithmetic Computation	1.1	1.1
Arithmetic Concepts	1.4	1.2
Arithmetic Applications	1.1	1.0
Social Studies	1.1	1.0
Science	1.1	1.0

[a]See Table 51 for mean grade equivalent scores from baseline testing.

GENERAL RESULTS OF EXPLORATORY FIELD TESTING: PROBLEMS AND RECOMMENDATIONS

In the 1986-1987 field-testing program, the basic strategy employed was the control group method. Typically, pretests were administered to both P & I Units and controls in the fall, and the same or an alternate form

of the test to the setting. Evaluations of service-wide uses that compared to serve as comparisons between the two groups.

Several serious problems were encountered with this procedure. First of all, securing an appropriate control group for any given Unit proved to be extremely difficult. In the absence of random assignment of students to groups (which, especially after the groups are in different schools and of groups to Code-name B & I Units was located to schools on the basis of a number of factors, including decisions of school staff representatives, building principals, teachers, and the staff of Project WISDOM in the Research and Development Center), schools were used to serve to secure control groups which are closely matched to B & I Units on as many as possible of the relevant characteristics, including racial mix, and achievement, ages and maturity to it, and family background. Such matching must, in practice, involve many compromises with the ideal, and some serious mismatching will inevitably occur.

In fact, analysis of covariance, which seeks to differences in initial performance but not to differential rates of change in performance, can be carried out only with the understanding that some of the underlying assumptions are going to be violated. In short, there will be unavoidable contamination of such matches that is to say to make the minimum differences arising because of programs designed to improve the performance of students in the control schools, teacher characteristics and interests which are inescapably contaminated with groups, differences in school facilities, utilization of resources, unique and uncontrollable characteristics of the individual classrooms and students involved, unusual features of curricula, and many other factors not enumerated.

With all these problems in mind, the staff of Project WISDOM and representatives of the Jacksonville, Madison, and Racine school systems nevertheless decided that in this first year of field testing, the control group approach was an appropriate method for obtaining process data and for determining how best to proceed in the future.

A second strategy for field testing was used successfully to supplement the control-group approach. In many cases, grade equivalent scores (using published norms) were obtained for all and spring testing, and mean gains calculated for each Unit. These gains were then compared with expected gains based on the published norms and hence the relationship between test administration, although not entirely satisfactory (see VII, p.), this strategy at least made it possible to make some meaningful comparisons.

A more desirable method would involve selection of data on a group of students over a period of several years prior to their inclusion in the B & I Unit. A pattern of growth could then be determined, and the effectiveness of the Unit evaluated on the basis of any deviations from this pattern. Techniques of time-series analysis might appropriately be employed in this context.)

Even more desirable would be the establishment of objective (and subjective) criteria for the success of the B & I Unit, and the development of appropriate instruments for assessing the success of the Unit in meeting these criteria. Such a procedure would be a response to the effect of subject-centered curriculum and behavioral scientists, and would include both cognitive and affective behaviors of both teachers and students.

Experience in the 1966-1967 field-testing program indicated that the use of standardized achievement tests to evaluate B & I Units is not completely satisfactory. The nature of such tests as the Stanford Achievement Test with their orientation toward traditional curricula results in their being inadequate to situations in which innovative and developmental activities are emphasized. When a third-grade B & I Unit focuses on the processes of scientific inquiry, while the tests are oriented toward specific knowledge about patterns in scientific facts, the results are unfortunate. Indeed, it is surprising that students in some B & I Units performed as well as they did on standardized tests.

[5] Worthen et al., op. cit.

[illegible]

The Opinion Scale was administered to morning specialists and teachers of E.S.L. Units, teachers in classrooms participating in contacts for joint reading sessions, and building principals of schools containing E.S.L. Units and control classes. The Questionnaire is limited to morning specialists and teachers since no complete data were available for the principals. To obtain an adequately large sample, data from twelve new cycles with two Units and personnel in purposes of analysis.

Table 10 indicates the mean item scores and subscores for various divisions of the questionnaire. For each item, the statement which presented the most favorable alternative was scored with descending scores representing more favorable statements, and a score of 5 or 6 indicating the least favorable. Each subscore is a sum of scores for the items related to that area of measurement. Maximum possible scores for each subdivision are as follows: Contributions 10, Research 10, Development 10, Innovation 10, Effect on Teachers 10, Effect on Students 10, and Utilization 8.

In every case the scores for the E.S.L. Unit personnel are greater than those for the control teachers indicating the superiority of this approach according to these areas of measurement. Substantial differences were between Unit scores and instruction subscores for these two groups. The only subscore that E.S.L. instructional staff rated the value of Research a model instructional program, of involvement in research projects, and of team planning. They were also more satisfied with their Unit instructional program.

The data were analyzed using a Wilcoxon rank procedure. A sum of scores for each alternative was obtained and the percent of each group choosing each alternative was determined. Where substantial differences were found between the E.S.L. and control groups, Student behavior, achievement, motivation, and attitudes were reportedly better in E.S.L. classes than in control classrooms. Moreover, greater satisfaction with student behavior and motivation was expressed by E.S.L. Unit personnel than by control teachers.

Table 10

Mean Differences on Teacher Opinion Questionnaire

Group	Total Score	Contribution	Research	Development	Innovation	Effect on Teachers	Effect on Students	Utilization
Learning Specialists E.S.L.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
E.S.L. Teachers	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
E.S.L. Total	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Control Teachers	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible] were encountered in carrying out the field testing plans for the Circle of Grace, adequate baseline data were not available for

[illegible] and the Hawaii School, and the district administration of the control school and those of the Hawaiian school were taken at the beginning of the academic year. Nevertheless, examination of the baseline data on the two groups indicated that the Hawaiian children were significantly behind the control group in their academic performance.

[illegible] were administered to the Hawaiian children in order to familiarize them with the testing procedures. Additionally, an experimental group of parents were trained individually at home by the field leader. The attitudes of these parents toward the school were significantly better than those of control parents at the end of the year. Furthermore, the academic performance of children in the experimental group was superior.

The field test data for this first semester Hawaii school student performance was significantly better than that of control school students on three of the standard subtests. A section part of seven months was made in the five month interval between the administrations. Considering the fact that median subject performance in October of fifth grade was only 3.3, the above evidence seems more than reasonable. Apparently the learning rate of these students, which had been consistently below average, increased markedly in the course of the year. The gain in achievement between the average fifth grader and Hawaii School fifth graders narrowed rather than continued to widening trend.

Motivational procedures in mathematics, together with individualization, were evaluated in three similar experiments in the Hawaii School second-grade unit and in two units at Hoaglane Hill School. All experiments involved